

PEDECIBA Informática
Instituto de Computación – Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay

Tesis de Maestría

en Informática

**Hacia una taxonomía de educación
de ingeniería de software**

Sebastián Pizard

2016

Hacia una taxonomía de educación de ingeniería de software
Sebastián Pizard
ISSN 0797-6410
Tesis de Maestría en Informática
Reporte Técnico RT 16-11
PEDECIBA
Instituto de Computación – Facultad de Ingeniería
Universidad de la República.
Montevideo, Uruguay, 2016

HACIA UNA TAXONOMÍA SOBRE EDUCACIÓN DE INGENIERÍA DE SOFTWARE

Sebastián Pizard
Tesis de Maestría en Informática
PEDECIBA

Tutor: Diego Vallespir

Facultad de Ingeniería
Universidad de la República
Montevideo, Uruguay / 2016

Insectes: 1. *Platypria echidna* (grossi). — 2. *Sagra* *Buqueti*. — 3. *Desmonota variolosa*. — 4. *Pachicleris* *Hawdicki*. — 5. *Catacanthus anehorago*. — 6. *Callidea signata*. — 7. *Eulyes melanoptera*. — 8. *Holco-
meria spinosa*. — 9. *Antioecis bilineatus*. — 10. *Cheirochela* *Fenn*. — 11. *Huechys incornata*. — 12. *Gosana frivata*. — 13. *Aphana novemmaculata*. — 14. *Hottius Delesserti*. — 15. *Phronia rubra*. —
16. *Tettigonia pulchella* (grossi). — 17. *Tettigonia colorata* (grossi). — 18. *Tettigonia fasciata* (grossi). — 19. *Phyllomorpha albigera*. — 20. *Philocerus fuscus* (grossi). — 21. *Carcinocoris Castetii* (grossi).
— 22. *Philonothobla* (grossi). — 23. *Hemiptycha* (grossi). — 24. *Spongophorus ballista* (grossi). — 25. *Spongophorus Benneti* (grossi). — 26. *Boecidium globulare* (grossi). — 27. *Cyphonia furcata* (grossi).
— 28. *Cyphonia tritida* (grossi). — 29. *Heteronotus vulnerans* (femelle) (grossi). — 30. *Heteronotus vulnerans* (male) (grossi). — 31. *Eda indata* (grossi). — 32. *Membracis C. album* (grossi). — 33. *Combophora*
Beakii (grossi). — 34. *Acanthopus splendidus*. — 35. *Xylocopa caerulea*. — 36. *Scolia procer*. — 37. *Spheg* *Latreille*. — 38. *Synagris cornuta*. — 39. *Eumenes arcuata*. — 40. *Pepais nureipennis*. — 41. *Mutilla*
cephalotes. — 42. *Chlorion lobatum*. — 43. *Pelocinus polycerator*. — 44. *Megiscus ruficeps*. — 45. *Chrysis semicincta* (grossi). — 46. *Cleptes orientalis* (grossi). — 47. *Euchromus purpuratus* (grossi). —
48. *Silphium salina* (grossi). — 49. *Anthrenus* *Stenopus*. — 50. *Mutilla* *clavicornis*. — 51. *Mutilla* *clavicornis*. — 52. *Mutilla* *clavicornis*. — 53. *Mutilla* *clavicornis*. — 54. *Mutilla* *clavicornis*. — 55. *Mutilla* *clavicornis*. — 56. *Mutilla* *clavicornis*. — 57. *Mutilla* *clavicornis*. — 58. *Mutilla* *clavicornis*. — 59. *Mutilla* *clavicornis*. — 60. *Mutilla* *clavicornis*. — 61. *Mutilla* *clavicornis*. — 62. *Mutilla* *clavicornis*. — 63. *Mutilla* *clavicornis*. — 64. *Mutilla* *clavicornis*. — 65. *Mutilla* *clavicornis*. — 66. *Mutilla* *clavicornis*. — 67. *Mutilla* *clavicornis*. — 68. *Mutilla* *clavicornis*. — 69. *Mutilla* *clavicornis*. — 70. *Mutilla* *clavicornis*. — 71. *Mutilla* *clavicornis*. — 72. *Mutilla* *clavicornis*. — 73. *Mutilla* *clavicornis*. — 74. *Mutilla* *clavicornis*. — 75. *Mutilla* *clavicornis*. — 76. *Mutilla* *clavicornis*. — 77. *Mutilla* *clavicornis*. — 78. *Mutilla* *clavicornis*. — 79. *Mutilla* *clavicornis*. — 80. *Mutilla* *clavicornis*. — 81. *Mutilla* *clavicornis*. — 82. *Mutilla* *clavicornis*. — 83. *Mutilla* *clavicornis*. — 84. *Mutilla* *clavicornis*. — 85. *Mutilla* *clavicornis*. — 86. *Mutilla* *clavicornis*. — 87. *Mutilla* *clavicornis*. — 88. *Mutilla* *clavicornis*. — 89. *Mutilla* *clavicornis*. — 90. *Mutilla* *clavicornis*. — 91. *Mutilla* *clavicornis*. — 92. *Mutilla* *clavicornis*. — 93. *Mutilla* *clavicornis*. — 94. *Mutilla* *clavicornis*. — 95. *Mutilla* *clavicornis*. — 96. *Mutilla* *clavicornis*. — 97. *Mutilla* *clavicornis*. — 98. *Mutilla* *clavicornis*. — 99. *Mutilla* *clavicornis*. — 100. *Mutilla* *clavicornis*.

RESUMEN

La educación de la ingeniería de software es una disciplina relativamente nueva y es objeto de amplia investigación. La literatura sobre el área de la educación de ingeniería de software consiste más que nada en artículos publicados en revistas y conferencias especializadas. Este material no se encuentra clasificado en categorías estables ni cuenta con una terminología estándar.

La importancia de organizar el conocimiento y las publicaciones que se realizan está asociada a los beneficios que se obtienen para diferentes involucrados, como son: permitir a los investigadores situar sus iniciativas en un contexto adecuado, lograr una mejor inducción a los términos y conceptos de un área temática, y además mejorar sustancialmente la acumulación de evidencia (revisiones sistemáticas de literatura y estudios de mapeo).

Este trabajo pretende contribuir a la organización del conocimiento de la educación de la ingeniería de software por medio de la construcción de un vocabulario controlado sobre los términos involucrados en la disciplina. Un vocabulario controlado es un conjunto de términos previamente seleccionados y sujeto a cierto control para su modificación. Estos términos son luego utilizados para describir documentos u otros tipos de objetos de contenido.

Se propone un proceso de construcción de vocabularios controlados basado en la exploración del material a clasificar del cual se realizaron dos instancias de ejecución. En una primera se utilizaron técnicas de *clustering* automático para analizar más de 1.000 artículos presentados en conferencias relevantes durante el período (1988-2014) y obtener una taxonomía inicial. En una segunda instancia se amplió la cantidad de términos de la categoría '*teaching approaches and methods*'.

El resultado es un vocabulario controlado sobre educación de ingeniería de software con forma de taxonomía cuyo nivel más alto lo componen tres facetas: '*what to teach*' (24 términos), '*how to teach*' (30 términos) y '*where to teach*' (6 términos). Los 26 términos de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' cuentan con una definición, basada en la literatura disponible, y una lista de referencias bibliográficas.

*How can one know that a particular chunk of metal is gold unless one knows what gold is, and how can one know what gold is without inspecting some samples of gold?
But if one does not know what gold is, one cannot decide what to inspect...*

— [Craig and Routledge (Firm), 1998] David L. Hull

AGRADECIMIENTOS

La elaboración de esta tesis fue un gran desafío personal. Cumplir este desafío sólo fue posible gracias a las personas que me acompañaron durante todo el proceso. A todos ellos les agradezco desde el corazón.

A mi tutor, Diego, por acompañarme en todo el proceso, por saber respetar mis tiempos de aprendizaje y exigirme en los momentos necesarios.

A toda mi familia, en especial a mi madre y mi padre los cuales han fomentado en mí las ganas de indagar e investigar el universo. También les tengo que agradecer su compañía en la vida y apoyo en todas las decisiones que he tomado.

A mis compañeros de facultad, a mis compañeros del GrIS, y en particular a Silvana, Lucyla, Mónica y Joseline a quienes atormenté durante toda la tesis.

Por último, y no por eso menos importante, tengo que agradecerle a dos personas muy importantes en mi vida. A ambos tengo que agradecerles el amor. A Mariel, por su estar en momentos difíciles y delicados. A Matías, por estos casi dos años de camino juntos.

ÍNDICE GENERAL

1	INTRODUCCIÓN	1
1.1	Motivación	1
1.2	Objetivos	3
1.3	Trabajo realizado	4
1.4	Publicaciones	5
1.5	Estructura del Documento	6
2	VOCABULARIOS CONTROLADOS	9
2.1	Introducción	9
2.2	Definición de vocabulario controlado	10
2.3	Conceptos básicos	12
2.4	Principios de los vocabularios controlados	13
2.4.1	Garantía Literaria	14
2.5	Estructuras de los vocabularios controlados	16
2.5.1	Listas	16
2.5.2	Anillos de sinónimos	17
2.5.3	Taxonomías	17
2.5.4	Tesauros	17
2.6	Construcción	18
2.7	Análisis de facetas	19
3	TRABAJOS RELACIONADOS	21
3.1	Introducción	21
3.2	Libros sobre Educación de Ingeniería de Software	21
3.3	Tópicos en las Conferencias sobre Educación de Ingeniería de Software	23
3.4	Revisión Terciaria sobre Educación de Ingeniería de Software	26
3.5	Vocabularios Controlados sobre Educación de Ingeniería	36
3.6	Conclusiones	38
4	MÉTODO DE TRABAJO Y APLICACIÓN	41
4.1	Proceso de construcción de vocabularios controlados	41
4.1.1	Definición de los objetivos y plan de trabajo	42
4.1.2	Recolección de vocabulario	44
4.1.3	Identificación de conceptos y relaciones	45
4.1.4	Validación del vocabulario controlado	46
4.2	Aplicación	46
5	VOCABULARIO CONTROLADO INICIAL	49

5.1	Método	49	
5.1.1	Definición de los objetivos y plan de trabajo	49	
5.1.2	Recolección de vocabulario	51	
5.1.3	Identificación de conceptos y relaciones	53	
5.1.4	Validación del vocabulario controlado	54	
5.2	Resultados	55	
5.2.1	What to teach	58	
5.2.2	How to teach	63	
5.2.3	Where to teach	64	
5.3	Limitaciones	66	
5.4	Conclusiones	66	
6	EXPANSIÓN DE LA CATEGORÍA ' <i>teaching approaches and methods</i> '		69
6.1	Método	69	
6.1.1	Definición de los objetivos y plan de trabajo	69	
6.1.2	Recolección de vocabulario	70	
6.1.3	Identificación de conceptos y relaciones	77	
6.1.4	Validación del vocabulario controlado	78	
6.2	Resultados	79	
6.3	Limitaciones	79	
6.4	Conclusiones	80	
7	CONCLUSIONES Y TRABAJOS A FUTURO		83
7.1	Conclusiones	83	
7.2	Trabajos a Futuro	88	
	BIBLIOGRAFÍA		93
A	TAXONOMÍA DE PALABRAS CLAVE PARA EER		99
A.1	Taxonomía de palabras clave para EER		99
A.2	Guía de uso de la taxonomía		108
A.2.1	Guidelines for assigning keywords		108
B	VOCABULARIO CONTROLADO INICIAL		111
C	CATEGORÍA ' <i>teaching approaches and methods</i> ' EXPANDIDA		115

ÍNDICE DE FIGURAS

Figura 1	Ilustraciones de Ehret, en la 5a. edición de <i>Systema Naturae</i> publicada en 1754, que muestra los caracteres utilizados para determinar las 24 clases de Linneo de plantas. 10	
Figura 2	Artículo con la presentación de la guía taxonómica de plantas que se comercializan como medicinales en Uruguay 13	
Figura 3	Incremento de la complejidad estructural en los vocabularios controlados. 16	
Figura 4	Ejemplo de taxonomía. 17	
Figura 5	Ejemplo de una entrada impresa de un tesaurus. 18	
Figura 6	Proceso para la construcción de un vocabulario controlado con un enfoque empírico. 42	
Figura 7	Proceso de creación del vocabulario controlado inicial sobre educación de ingeniería de software. 50	
Figura 8	Taxonomía inicial sobre educación de ingeniería de software obtenida en la primera instancia. 57	
Figura 9	Evolución de los conceptos de la faceta ' <i>what to teach</i> ' a lo largo del tiempo. 58	
Figura 10	Evolución de los conceptos de la faceta ' <i>how to teach</i> ' a lo largo del tiempo. 63	
Figura 11	Evolución de los conceptos de la faceta ' <i>where to teach</i> ' a lo largo del tiempo. 65	
Figura 12	Proceso de expansión de la categoría ' <i>teaching approaches and methods</i> ' de la faceta ' <i>how to teach</i> ' del vocabulario controlado sobre educación de ingeniería de software. 71	
Figura 13	Ejemplo de información extraída para el concepto ' <i>Lectures</i> ' 76	
Figura 14	Proceso de síntesis temática adaptado de Creswell [2011] 77	
Figura 15	Ejemplo de códigos en el etiquetado de fragmentos de textos sobre el concepto ' <i>Lectures</i> ' 78	
Figura 16	Taxonomía inicial sobre educación de ingeniería de software obtenida en la segunda instancia. 82	

Figura 17	Taxonomía de palabras clave para EER - pág. 1	100
Figura 18	Taxonomía de palabras clave para EER - pág. 2	101
Figura 19	Taxonomía de palabras clave para EER - pág. 3	102
Figura 20	Taxonomía de palabras clave para EER - pág. 4	103
Figura 21	Taxonomía de palabras clave para EER - pág. 5	104
Figura 22	Taxonomía de palabras clave para EER - pág. 6	105
Figura 23	Taxonomía de palabras clave para EER - pág. 7	106
Figura 24	Taxonomía de palabras clave para EER - pág. 8	107
Figura 25	Taxonomía de palabras clave para EER - Guía para selección de palabras clave.	109

ÍNDICE DE CUADROS

Cuadro 1	Tópicos de los llamados a trabajos de las últimas ediciones de CSEET e ICSE. 23
Cuadro 2	Revisión terciaria sobre educación de ingeniería de software 27
Cuadro 3	Categorías del esquema de clasificación de Nascimento et al. [2013] 29
Cuadro 4	Categorías del esquema de clasificación de Marques et al. [2014] 31
Cuadro 5	Categorías sobre investigación de educación de ingeniería de software basadas en agrupaciones de Ellis et al. [2009] con descripciones agregadas por Malik and Zafar [2012]. 35
Cuadro 6	Corpus considerado en la construcción del vocabulario controlado. 55
Cuadro 7	Correspondencia entre los conceptos de la faceta <i>'what to teach'</i> y el temario del SWEBOK. 60
Cuadro 8	Correspondencia entre los conceptos de la categoría <i>'teaching approaches and methods'</i> y el esquema de clasificación de Nascimento et al. [2013]. 64
Cuadro 9	Lista ampliada de conceptos candidatos para incluir en la categoría <i>'teaching approaches and methods'</i> 72
Cuadro 10	Ejemplo de cadena de búsqueda para el concepto <i>'Lectures'</i> . 74
Cuadro 11	Entrada de ejemplo con el término <i>Lectures</i> 79
Cuadro 12	Vocabulario controlado inicial sobre Educación de ingeniería de software 111
Cuadro 13	Categoría <i>'teaching approaches and methods'</i> de la faceta <i>'how to teach'</i> expandida. 115

INTRODUCCIÓN

El rol cada vez más crítico del software presenta nuevos desafíos para la educación de los ingenieros de software. No sólo nuestra dependencia del software ha aumentado sino que también ha cambiado la naturaleza de la producción del software y con eso las demandas sobre quienes lo desarrollan [Shaw, 2000].

Este trabajo pretende contribuir a la organización del conocimiento de la educación de la ingeniería de software y para esto busca construir un vocabulario controlado sobre los términos involucrados en la disciplina. Un vocabulario controlado es un conjunto de términos previamente seleccionados y sujeto a cierto control para su modificación. Estos términos son utilizados para describir documentos u otros tipos de objetos de contenido, logrando así organizar el material según los elementos que se hayan elegido para describirlo [NISO,2005].

Este capítulo se estructura de la siguiente forma. En la sección 1.1 se incluye la motivación. Más adelante, en la sección 1.2, se presentan los objetivos de esta tesis. Un resumen del trabajo realizado se incluye en la sección 1.3. En la sección 1.4 se listan las publicaciones realizadas en el marco de la tesis. Por último, en la sección 1.5 se describe la estructura de este informe.

1.1 MOTIVACIÓN

La educación de la ingeniería de software es una disciplina relativamente nueva y es objeto de amplia investigación. Comprende una gran variedad de tópicos como son: habilidades de comunicación, resolución de problemas, metodologías de diseño, negociación, interacción persona-computadora, liderazgo, ética, entre otras [Freeman et al., 1976; Lethbridge, 2000; Bourque and Fairley, 2014]. Actualmente se publica investigación sobre educación de ingeniería de software en distintos lugares. Estos lugares de publicación son, en general, revistas o conferencias y tienen objetivos diferentes: mientras que algunos son más específicos otros son más generales, y comprenden, por ejemplo, educación de ingeniería o educación en su definición más vasta. En definitiva, se publica una gran cantidad de artículos científicos en distintos lugares con distintos objetivos y con distintos niveles de abstracción temática (educación genérica, educación de ingeniería, educación de ingeniería de software).

Para recuperar la información se utilizan palabras clave, las cuales son asignadas por los autores y/o editores previo a la publicación de cada artículo. En la práctica, sucede que la calidad de la información recuperada por los investigadores depende de la calidad de las palabras clave utilizadas para etiquetar los artículos. En áreas poco maduras o poco estudiadas, como la educación de la ingeniería de software, no existen listas de palabras clave consensuadas. Esto dificulta tanto el etiquetado de los artículos como la recuperación de la información.

Estandarizar terminología en un área temática permite contar con beneficios para múltiples involucrados [Finelli et al., 2015]: puede servir como plataforma para que investigadores y miembros de la comunidad sitúen sus iniciativas de investigación en un contexto más amplio; puede ser una fuente de términos para entender la amplitud del campo y conducir búsquedas en la literatura; puede ser un mecanismo para que editores de revistas científicas puedan organizar autores afines y áreas de investigación, identificar la adecuación de los artículos enviados a la misión de la revista y categorizar la base de datos de revisores para una asignación más apropiada; y puede proveer de un método para que agencias de financiamiento clasifiquen mejor las áreas de investigación para lograr una mejor distribución de los fondos.

Otro beneficio relevante está asociado a la mejora de la acumulación de evidencia, por ejemplo, mediante las revisiones sistemáticas de literatura y los estudios de mapeo. Kitchenham et al. resumen la dificultad actual de no contar con una terminología estándar a utilizar en las revisiones sistemáticas de literatura para toda el área la ingeniería de software:

"Software engineering lacks strong taxonomies. The terms that we use are often imprecise, and software engineers are rather prone to create new terms to describe ideas that may well be closely related to existing ones. This can complicate searching since we need to consider all possible forms of terminology that might have been used in the titles and abstracts of papers. Snowballing may help with this, but essentially it stems from the constraint of studying procedures and artifacts" [Kitchenham et al., 2015, p. 26].

La estandarización de la terminología ha sido utilizada en áreas de la educación. Por ejemplo, Finelli et al. [2015] realizaron un trabajo de investigación muy importante para construir una taxonomía de palabras clave (*keywords*) para la investigación de educación de ingeniería. Aunque la taxonomía hace foco en la investigación de educación de ingeniería de Estados Unidos, los autores trataron de que fuera un esfuerzo inclusivo logrando la colaboración de 266 individuos de 30 países durante *workshops*, sesiones de conferencias y otras ac-

tividades. El resultado comprende una taxonomía de 455 términos organizados en 14 ramas de seis niveles.

También este enfoque se ha utilizado en otras ramas de la ciencia, por ejemplo, Michie et al. [2015] presentan un proyecto de 3 años con centenares de colaboradores de distintos países en el cual han desarrollado y evaluado una taxonomía de técnicas de cambio de comportamiento (usadas para cambiar hábitos como el consumo de drogas o alcohol). Según los investigadores que guiaron el trabajo, las intervenciones con técnicas de cambio de comportamiento eran reportadas utilizando diferentes etiquetas y muchas veces se utilizaba una misma etiqueta para diferentes técnicas. La falta de definiciones estándar significaba que cada investigador que realizaba revisiones sistemáticas utilizaba sus propios sistemas de clasificación, esto llevaba en muchos casos a duplicar esfuerzos y debilitar la posibilidad de acumular evidencia entre revisiones. Como resultado obtuvieron una taxonomía lograda a partir del consenso de investigadores que cubre 93 técnicas de cambio de comportamiento con etiquetas claras, definiciones y ejemplos.

No hemos encontrado iniciativas de organización del conocimiento o análisis de los artículos de investigación publicados en el área específica de la educación de ingeniería de software. Creemos que es necesario y oportuno comenzar a trabajar en la organización del conocimiento y la estandarización de la terminología utilizada en esta área temática.

1.2 OBJETIVOS

El objetivo general en el cual se enmarca este trabajo es contribuir en la organización del conocimiento sobre el área temática de la educación de la ingeniería de software.

El objetivo específico es la creación de un vocabulario controlado sobre educación de ingeniería de software. La construcción de un vocabulario controlado, visto como un esquema de clasificación, es un enfoque empírico posible para entender un área temática [Hjørland, 2002]. Los vocabularios controlados pueden tener distintas formas: desde listas de términos hasta taxonomías o tesauros. Permiten ordenar el conocimiento contemplado en un área temática mediante una lista unificada o categorización de los tópicos que comprende.

Se plantean las siguientes actividades para buscar alcanzar el objetivo específico:

1. Estudiar conceptos teóricos sobre organización del conocimiento mediante vocabularios controlados y taxonomías; y en particular, relevar guías y estándares para su construcción.

2. Relevar estudios y trabajos sobre la creación y uso de vocabularios controlados y esquemas de clasificación en general sobre educación de ingeniería de software.
3. Proponer un proceso para la creación de un vocabulario controlado que cumpla con guías y estándares reconocidos del área de organización del conocimiento.
4. Utilizar el proceso propuesto para elaborar un vocabulario controlado inicial sobre educación de ingeniería de software.

1.3 TRABAJO REALIZADO

Para realizar este trabajo fue necesario entender los conceptos y fundamentos de la organización del conocimiento mediante taxonomías y vocabularios controlados. Para esto se estudiaron: la guía Z39.19-2005 provista por ANSI/NISO [NISO,2005] y los libros de Hedden [2010] y Lambe [2007].

Ante la falta de un proceso con el nivel de formalidad adecuado para utilizar en este trabajo se definió un proceso disciplinado que guía en la creación o actualización de un vocabulario controlado utilizando el material a clasificar como base para la recolección de conceptos a considerar. Este proceso se basa en las guías y recomendaciones antes descritas.

El proceso propuesto cuenta con 4 etapas bien marcadas: definición de objetivos, recolección de vocabulario, identificación de conceptos y relaciones; y validación del vocabulario controlado. Es genérico y permite que diferentes técnicas o métodos puedan ser utilizadas para las distintas etapas.

Este proceso fue instanciado para construir un primer vocabulario controlado sobre educación de ingeniería de software. Para obtener una lista de los conceptos más investigados en educación de ingeniería de software se tomaron como base los artículos publicados en la *Conference on Software Engineering Education and Training (CSEET)* durante el período 1988-2014 y en el *Software Engineering Education and Training Track (SEET)* de la *International Conference on Software Engineering (ICSE)* durante el período 2000-2014 (sumando un total de 1.023 artículos). Con el objetivo de estudiar un corpus de tantos artículos y extraer posibles conceptos a incluir en el vocabulario se utilizó una herramienta de *clustering* automático que permitió analizar los resúmenes de los artículos en forma automática y detectar conjuntos de artículos similares. El resultado es un vocabulario controlado en forma de taxonomía con 43 términos que corresponden a los conceptos que se identificaron como los más utilizados en las publicaciones de investigación en el período cubierto. Estos términos se clasi-

ficaron en tres grandes facetas: *'what to teach'*, *'how to teach'* y *'where to teach'*. Se estudió también la evolución de los conceptos encontrados a lo largo de los años de publicación.

Se identificó la necesidad de ampliar la faceta *'how to teach'* pues no se encontraron referencias directas de los conceptos que abarca. Para esto se instanció nuevamente el proceso de creación de vocabularios controlados para una primera expansión de la categoría *'teaching approaches and methods'* de la faceta *'how to teach'* del vocabulario controlado construido anteriormente. Los objetivos específicos de esta instancia fueron ampliar la cantidad de conceptos de la categoría e identificar una definición concisa y lo menos ambigua posible para cada uno. Fue necesario adaptar dos métodos: uno para la recolección de vocabulario en detalle y otro para la síntesis de todo el material de un concepto en una única definición. Mediante ambos métodos fue posible estudiar más de 250 artículos sobre 34 conceptos candidatos y se obtuvieron 26 términos con sus definiciones y referencias más relevantes.

Luego de la ejecución de ambas instancias del proceso se obtuvo un vocabulario controlado sobre educación de ingeniería de software con forma de taxonomía cuyo nivel más alto lo componen tres facetas: *'what to teach'* (24 términos), *'how to teach'* (30 términos) y *'where to teach'* (6 términos). Los 26 términos de la categoría *'teaching approaches and methods'* de la faceta *'how to teach'* cuentan con una definición, basada en la literatura disponible, y una lista de referencias bibliográficas.

1.4 PUBLICACIONES

En el contexto de esta tesis un artículo que resume el trabajo presentado en el capítulo 5 fue aceptado para su publicación en el *European Journal of Engineering Education (EJEE)*.

La EJEE es la revista científica oficial de la Asociación Europea para la Educación de Ingeniería (*European Society for Engineering Education - SEFI*). Tiene una publicación de seis números por año. Se publica desde 1976 y está en plena campaña para aparecer en el JCR. Actualmente su índice SJR es de 0.856¹.

Pizard, S. and Vallespir, D. "Towards a Controlled Vocabulary on Software Engineering Education," *European Journal for Engineering Education*, 2016.

¹ <http://www.scimagojr.com/journalsearch.php?q=11300153314&tip=sid&clean=o>

El artículo aceptado presenta: la motivación para realizar un vocabulario controlado sobre educación de ingeniería de software, el método propuesto y los resultados obtenidos en su primera instancia (reportada en el capítulo 5).

1.5 ESTRUCTURA DEL DOCUMENTO

El informe de esta tesis cuenta con 6 capítulos además del actual. El capítulo 2 "*Vocabularios Controlados*" presenta los principales conceptos asociados a los vocabularios controlados, incluyendo su definición y formas más utilizadas. Este capítulo se basa en la guía Z39.19-2005 [NISO,2005] y el libro de Hedden [2010].

En el capítulo 3 "*Trabajos Relacionados*" se presentan los resultados del relevamiento de estudios previos sobre creación y uso de vocabularios controlados y esquemas de clasificación en general sobre educación de ingeniería de software.

Luego, en el capítulo 4 "*Método de Trabajo y Aplicación*" se describe el proceso propuesto para la construcción de vocabularios controlados y se indica el plan de su aplicación en el área de la educación de ingeniería de software.

La primera instancia de aplicación del proceso para la creación de una taxonomía inicial sobre educación de ingeniería de software se describe en el capítulo 5 "*Vocabulario Controlado Inicial*". Allí se presentan los objetivos del trabajo, las técnicas utilizadas para la recolección del vocabulario, las limitaciones del método utilizado así como los resultados obtenidos. En el capítulo 6 "*Expansión de la Categoría 'teaching approaches and methods'*" se describe una segunda aplicación del proceso, en este caso para realizar una primera expansión de la categoría '*teaching approaches and methods*' (enfoques y métodos de enseñanza) de la faceta '*how to teach*'. La expansión comprende ampliar el alcance de esa categoría e identificar una definición concisa y lo menos ambigua para cada uno de sus términos.

Por último, el capítulo 7 "*Conclusiones y Trabajos a Futuro*" presenta reflexiones sobre el trabajo realizado y posibles tareas a realizar en futuros trabajos.

El informe contiene también tres Anexos. En el Anexo A "*Taxonomía de palabras clave para investigación sobre educación de ingeniería*" se presenta la taxonomía de palabras clave (*keywords*) para la investigación sobre educación de ingeniería (EER, por sus siglas en inglés) creada por Finelli et al. [2015], esto corresponde a un trabajo previo que se describe en la Sección 3.5. En el Anexo B "*Vocabulario Controlado Inicial sobre Educación de Ingeniería de Software*" se presenta el vocabulario obtenido en la primera instancia del trabajo mientras que en el Anexo C "*Categoría 'teaching approaches and methods' expandida*" se presenta la compo-

sición de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' obtenida durante la instancia de trabajo de expansión.

En este capítulo se presentan los principales conceptos asociados a los vocabularios controlados, incluyendo su definición e importancia. También se enumeran los principios que guían el control del vocabulario y los distintos tipos de vocabularios controlados según su estructura. Por último, se tratan las estrategias y técnicas para la construcción de vocabularios controlados. Este capítulo se basa en la la guía Z39.19-2005 provista por ANSI/NISO [NISO,2005] y las recomendaciones de Hedden [2010].

2.1 INTRODUCCIÓN

Las personas clasificamos objetos de la realidad regularmente. Si llueve decidimos qué tipo de zapatos debemos utilizar y al volver del supermercado ordenamos las comida comprada dependiendo de si necesita refrigeración o no. La mayoría del tiempo clasificamos sin darnos cuenta que lo estamos haciendo.

El propósito de la clasificación es organizar el conocimiento. La clasificación permite la abstracción de objetos individuales agrupándolos, según sus características, en distintas clases. La ciencia que trata los principios, métodos y fines de la clasificación es la taxonomía; aunque es más común utilizar esa palabra para referirse a un esquema de clasificación particular.

Una de las primeras taxonomías modernas fue la taxonomía linneana que clasifica a los seres vivos en diferentes niveles jerárquicos. En el siglo XVIII Carl Linnaeus investigó los seres vivos y creo un esquema que los clasificaba según características que eran observables en ese momento; como ejemplo se muestra en la Figura 1 su división del reino vegetal en 24 clases, cada una de las cuales se denominan de acuerdo con el número de estambres y su disposición en las flores. Más tarde el uso de taxonomías se propagó a otras ciencias y se utilizaron también para clasificar conceptos (una de las más conocidas es la taxonomía de objetivos educacionales o taxonomía de Bloom [1956]).

El término taxonomía presenta dos usos comunes. Un significado, que refleja su uso temprano en la clasificación de organismos vivos, es la clasificación jerárquica de cosas o conceptos. Los términos de la taxonomía tienen un «padre» (un término más amplio) y uno o más «hijos» (términos más específicos).



Figura 1: Ilustraciones de Ehret, en la 5a. edición de *Systema Naturae* publicada en 1754, que muestra los caracteres utilizados para determinar las 24 clases de Linneo de plantas.

Según Hedden [2010], otro uso más reciente del término taxonomía refiere a cualquier vocabulario controlado de términos de un área de dominio temática o para un propósito específico. Los términos pueden no estar organizados en una jerarquía y pueden o no tener relaciones más complejas entre ellos.

2.2 DEFINICIÓN DE VOCABULARIO CONTROLADO

Un vocabulario controlado es un conjunto de términos previamente seleccionados y sujeto a cierto control para su modificación. El propósito de los vocabularios controlados es proveer una manera de organizar la información. Esto

se logra mediante el proceso de asignar términos seleccionados del vocabulario controlado para describir documentos u otros tipos de objetos de contenido, logrando así organizar el material según los elementos que se hayan elegido para describirlo [NISO,2005]. Como ejemplo, el extracto de la taxonomía lineana presentada en la Figura 1 permite asignar una clase para describir cada vegetal. De esta forma, los vegetales quedan organizados según la cantidad y disposición de sus estambres (elementos elegidos para describirlos).

La necesidad de un vocabulario controlado surge de dos características básicas del lenguaje:

- Dos o más palabras o términos pueden ser usadas para representar un mismo concepto (i.e. sinónimos).
- Dos o más palabras o términos que se escriben de igual forma representan conceptos diferentes (i.e. homógrafos).

También surge de la necesidad de establecer relaciones entre los términos, por ejemplo, de relaciones de jerarquía para el caso de vocabularios controlados con estructura taxonómica.

El control del vocabulario se lleva a cabo mediante tres métodos fundamentales:

- Definir alcance, o significado, de los términos.
- Utilizar la relación de equivalencia para vincular términos sinónimos o casi sinónimos.
- Distinguir entre términos homógrafos (aquellas palabras que se escriben de forma idéntica pero tienen diferentes significados).

Se puede ver, por tanto, que un vocabulario controlado busca asegurar consistencia en la aplicación del lenguaje [Hedden, 2010]. Se dice que es controlado porque únicamente los términos de la lista deberían ser utilizados para referirse a los conceptos del área temática cubierta por el vocabulario. También es controlado ya que la modificación de la lista está sujeta a políticas definidas previamente.

En la actualidad el control del vocabulario es utilizado para mejorar la eficacia de los sistemas de almacenado y recuperación de información, sistemas de navegación web, y otros tipos de ambientes que buscan identificar y localizar contenido mediante algún tipo de descripción utilizando lenguaje [Hedden, 2010].

2.3 CONCEPTOS BÁSICOS

Un objeto de contenido es cualquier elemento que debe ser descrito para su inclusión en un sistema de recuperación de información, sitio web, u otra fuente de información. Objetos de contenido típicos son artículos de revistas científicas, informes técnicos, y otros tipos de documentos. Los objetos de contenido pueden existir en forma electrónica y también pueden ser entidades como pinturas, muebles o edificios.

Un concepto es una cosa, una idea, o un entendimiento compartido sobre algo; por otro lado, un término es una etiqueta para un concepto, en general la más común, una designación genérica.

Entonces, más formalmente, *un vocabulario controlado consiste en un conjunto de términos que han sido enumerados explícitamente para nombrar a los conceptos que han sido seleccionados para describir a objetos de contenido.*

En general, un vocabulario controlado es construido y actualizado por cierta organización para ser utilizado por ciertos usuarios. Como ejemplo, la guía taxonómica de plantas que se comercializan como medicinales en Uruguay (presentada en la Figura 2) fue creada por la Cátedra de Botánica Farmacéutica de la Facultad de Química para ser utilizada por médicos, químicos, farmacéuticos, y herboristas. Muchas características de un vocabulario controlado dependen de la organización y de los usuarios. En el ejemplo anterior, en la taxonomía de plantas seguramente se utilice vocabulario más técnico que el que se usaría si estuviera construido para la población en general.

Existen varias actividades relacionadas a la clasificación de objetos de contenido.

ETIQUETADO (TAGGING): A veces llamado *keywording*, involucra la creación de nuevos términos a medida que se revisan los objetos de contenido; en este caso, no se usa un vocabulario controlado y por tanto hay poco o ningún control de autoridad. En general, los resultados del etiquetado de un mismo conjunto de objetos de contenido realizado por distintas personas son diferentes y dependen de sus habilidades, conocimientos y rigurosidad del proceso llevado a cabo.

CATALOGACIÓN: Usualmente involucra la asignación a objetos de contenido de términos descriptivos seleccionados de un vocabulario controlado (catalogación de temas) pero también incluye la inclusión de otro tipo de metadata que puede ser bibliográfica, histórica o físico-descriptiva (catalogación descriptiva).

Acta Farm. Bonaerense 8 (3): 187-98 (1989)
 Recibido el 23 de junio de 1989

Actualizaciones

Guía Taxonómica de Plantas que se Comercializan como Medicinales en Uruguay. Primera Contribución.

HELENA PRIORE, LUCIA CARPIUC,
 EDUARDO ALONSO y MARÍA JULIA BASSAGODA
*Cátedra de Botánica Farmacéutica, Facultad de Química,
 Montevideo, Uruguay. Casilla de Correo 1157*

RESUMEN. Se presenta una lista de 175 especies de plantas que se comercializan como medicinales a nivel de herboristerías en la ciudad de Montevideo (Uruguay). Las muestras se obtuvieron en el comercio y se determinaron taxonómicamente. En la lista figuran ordenadas alfabéticamente las especies determinadas, familia, sinonimia, nombre vulgar y partes comercializadas.
 SUMMARY. "Taxonomic Guide of Plants Sold as Medicinal in Uruguay. First Contribution". A list of 175 plants, sold at herbalist shops in Montevideo (Uruguay) as having medicinal properties is presented. Samples were purchased and taxonomically determined. The list shows in alphabetical order, the determined species, family, synonymy, common name and the parts which are comercialized.

Figura 2: Artículo con la presentación de la guía taxonómica de plantas que se comercializan como medicinales en Uruguay

CLASIFICACIÓN: Comprende el proceso de asignar un ítem a una única clase de una lista de clases predeterminadas. Asignar nombres de organismos en una taxonomía biológica o nombrando conceptos técnicos o científicos y agrupándolos también se consideran formas de clasificar.

INDEXACIÓN (INDIZACIÓN): Implica crear un índice, una lista de términos, cada uno de los cuales apunta a la ubicación de la información o contenido del tema deseado. La indexación utilizando un vocabulario controlado implica mantener la lista de objetos de contenido para cada término. Esto es lo que realizan muchas plataformas de recuperación de información.

2.4 PRINCIPIOS DE LOS VOCABULARIOS CONTROLADOS

Hay cuatro principios importantes que guían el diseño y la construcción de un vocabulario controlado:

- *Eliminar la ambigüedad* - En el lenguaje natural ocurre ambigüedad cuando una palabra o frase tiene más de un significado. Por ejemplo, el término Mercurio puede referirse al planeta o al metal. Un vocabulario controlado debe compensar los problemas de ambigüedad asegurando que cada término tenga un único significado.
- *Controlar los sinónimos* - La sinonimia ocurre cuando un concepto puede ser representado por dos o más términos diferentes. La existencia de términos equivalentes para clasificar objetos de contenido puede causar dispersión: objetos que deberían estar agrupados juntos quedan en diferentes grupos según el término utilizado en su clasificación. Un vocabulario controlado debe compensar los problemas causados por la sinonimia asegurando que cada concepto sea representado por un término preferido. Los términos equivalentes para un mismo concepto serán registrados en el vocabulario indicando que debe utilizarse el término preferido para representarlo.
- *Establecer relaciones entre los términos cuando sea apropiado* - Un vocabulario controlado debe modelar las relaciones entre los términos que sean necesarias. Es posible modelar relaciones de equivalencia (explicada en el punto anterior), relaciones jerárquicas (indicando que un concepto es más específico que su concepto padre) y relaciones asociativas (indicando otro tipo de relaciones entre dos conceptos).
- *Probar y validar los términos* - El proceso de construcción de un vocabulario controlado involucra consultar varias fuentes de palabras y frases así como criterios basados en: el lenguaje natural utilizado para describir objetos de contenido (garantía literaria), el lenguaje de los usuarios (garantía de los usuarios) y las necesidades de la organización (garantía organizacional).

2.4.1 *Garantía Literaria*

El principio de garantía literaria es muy relevante para el trabajo realizado en esta tesis. Esta subsección brinda una mejor introducción a este principio, es necesario aclarar que se basa fuertemente en la investigación realizada por Barité [2009, 2011].

La primera referencia al principio de garantía literaria es la brindada por Hulme:

“las clases [es decir, los términos de un vocabulario controlado que representan conceptos] y las relaciones en los sistemas de clasificación deben ser decididas sobre la base de la literatura que ha de ser clasificada” ([Hulme, 1911–1912] *apud* [Barité, 2009]).²

Este autor también acota lo siguiente:

“Literary warrant meaning that the basis for classification is to be found in the actual published literature rather than abstract philosophical ideas or concepts in the universe of knowledge or the order of nature and system of the sciences” ([Hulme, 1911–1912] *apud* [Barité, 2011]).³

En la práctica esto significa que es la misma documentación la cual valida los términos a ser incluidos en un sistema de organización del conocimiento, ya que lo que se clasifica e indiza son temas presentes en los documentos [Barité, 2011].

Algunas de las razones por las cuales conviene que los sistemas de organización del conocimiento deberían apoyarse antes en la garantía literaria que en las clasificaciones científicas o especializadas son las siguientes [Barité, 2011]:

- En primer lugar, utilizar la garantía literaria tiene un beneficio más bien pragmático pues la función típica de la clasificación es la organización lógica de un conjunto de documentos. Al utilizar los mismos documentos para crear el sistema de clasificación es natural que ese sistema los permita clasificar.
- La documentación presenta habitualmente los procesos completos de construcción del conocimiento mientras que las clasificaciones científicas y especializadas sólo muestran los acuerdos actuales.
- Mientras que las clasificaciones científicas y especializadas incluyen únicamente las relaciones paradigmáticas entre conceptos (aquellas esenciales y generalmente estables que se establecen entre ideas, objetos y/o fenómenos) los sistemas de organización del conocimiento necesitan representar tanto las relaciones paradigmáticas como otro tipo de relaciones.
- A través de la garantía literaria, entre otros mecanismos, se facilita los procesos de desambiguación y relación entre términos. Esto es muy necesario en los sistemas de organización del conocimiento y muchas veces no es algo prioritario en la clasificaciones científicas o especializadas.

² La cita fue reproducida tal cual está en el original.

³ La cita fue reproducida tal cual está en el original.

De igual manera, no debe considerarse a la garantía literaria como opuesto a las clasificaciones científicas. Es más, parece sano combinar técnicas de análisis de expertos y garantía literaria para la construcción de sistemas de organización del conocimiento y vocabularios controlados en general.

Sobre si la garantía literaria debe extenderse en el tiempo, o sólo a la documentación que podemos llamar actual, ([Tennis, 2002] *apud* [Barité, 2011]) opina que es necesario mantener una perspectiva histórica, por lo que el sistema resultante debe proporcionar acceso a viejas concepciones del universo de materiales publicados. Esto significa que se requiere entender a la garantía literaria como un principio dinámico, que es capaz de asumir, absorber y representar los cambios y los avances que se producen en el mundo del conocimiento ([Foskett, 1996] *apud* [Barité, 2011]).

2.5 ESTRUCTURAS DE LOS VOCABULARIOS CONTROLADOS

Existen varios tipos de vocabularios controlados que contemplan diferentes propósitos y diferentes complejidades estructurales. En la Figura 3 se muestran los diferentes tipos de vocabularios controlados y sus características [NI-SO,2005, p. 16].



Figura 3: Incremento de la complejidad estructural en los vocabularios controlados.

2.5.1 Listas

En su forma más simple un vocabulario controlado consta únicamente de una lista de términos (cada uno asociado a un concepto distinto). En las listas sólo interesa eliminar la ambigüedad, para esto se define un único término para referirse a un concepto. Diferentes términos describen diferentes conceptos. Un ejemplo es la siguiente lista de plantas: saúco, rosa, pino, romero, helechos, orquídeas.

2.5.2 Anillos de sinónimos

Aunque los anillos de sinónimos son considerados un tipo de vocabulario controlado, se utilizan de forma diferente a los demás. No se usan para clasificar objetos de contenido, sino para su recuperación. Un anillo de sinónimos es un conjunto de términos que son considerados equivalentes. Un anillo de sinónimos permite a los usuarios acceder a todos los objetos de contenido que contengan alguno de esos términos. Un ejemplo es el siguiente conjunto de términos: pinos, abetos, coníferas, ocote.

2.5.3 Taxonomías

Una taxonomía es un vocabulario controlado que consta de términos preferidos conectados en una jerarquía o polijerarquía. Es muy importante identificar las relaciones de jerarquía entre los términos e indicarlas de forma explícita. La Figura 4 presenta un ejemplo de un vocabulario controlado con este tipo de estructura.

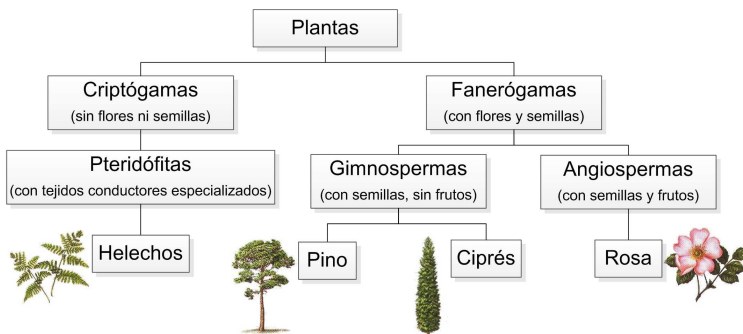


Figura 4: Ejemplo de taxonomía.

2.5.4 Tesauros

Por último, un tesoro es un vocabulario controlado arreglado según un orden conocido y estructurado para permitir mostrar claramente todos los tipos de relaciones que se requiere modelar. Esto significa que además de relaciones de jerarquía es posible representar cualquier otro tipo de relaciones entre los términos (por ejemplo: 'relacionado a', 'usa', 'dentro de', etc). La Figura 5 pre-

senta un ejemplo de una entrada de un tesoro. Los sinónimos (términos no preferidos) están indicados luego de la etiqueta UF (*used for*), los términos más generales en la jerarquía están indicados con BT (*broader term*) y los términos más específicos en la jerarquía están indicados con NT (*narrow term*).

Gimnospermas
 Son plantas vasculares y espermatofitas, productoras de semillas. Poseen semillas desnudas, sin fruto que las proteja. Son las plantas, con semillas, más antiguas.

UF Gymnospermae
 BT **Fanerógamas**
 NT **Pino, Ciprés**

Figura 5: Ejemplo de una entrada impresa de un tesoro.

2.6 CONSTRUCCIÓN

Para la construcción de un vocabulario controlado se plantean tres enfoques diferentes [NISO,2005, p. 91]:

- *Enfoque del comité* - Este enfoque implica que expertos del área redacten una lista con los términos clave indicando sus relaciones. Las listas de términos pueden ser presentadas por diversos expertos o tomadas de diversas fuentes. El enfoque del comité tiene dos métodos básicos:
 - *Top down* - Primero se identifican los términos más amplios y luego se seleccionan los términos más concretos según el nivel de especificidad elegido. La estructura jerárquica se crea durante el proceso de trabajo.
 - *Bottom up* - Igual que en el caso anterior la estructura jerárquica se construye durante el proceso de trabajo, pero se comienza con los términos más específicos primero y luego con los más amplios.

Este enfoque plantea como gran ventaja que son los expertos del área (que generalmente, serán también los futuros usuarios) los que construyen el vocabulario. Un ejemplo de aplicación de este enfoque es la construcción de la taxonomía de técnicas de cambio de comportamiento [Michie et al., 2015]. En este trabajo los autores ejecutaron un amplio conjunto de ejercicios de consenso utilizando el enfoque de expertos, primero mediante

métodos Delphi (por información sobre este método consultar el trabajo de Linstone et al. [1975]) para la recolección de las técnicas, luego aplicando métodos *top-down* y *bottom-up* para el desarrollo de la jerarquía de la taxonomía [Cane et al., 2015] y por último realizando actividades junto a los expertos para evaluar la taxonomía resultante.

- *Enfoque empírico* - Este enfoque contempla básicamente dos métodos.
 - *Método deductivo* - Los términos son extraídos de un conjunto de referencia de objetos de contenido (por humanos o computadoras), opcionalmente durante una etapa de indexación, pero no se intentan controlar ni determinar posibles relaciones hasta que un número suficiente de términos es recolectado. Luego, expertos (cuando sea posible) revisan los términos, identifican posibles jerarquías y asignan a cada una los términos restantes. El control del vocabulario se realiza a partir de entonces.
 - *Método inductivo* - Nuevos términos son seleccionados para su posible inclusión en el vocabulario a medida que son encontrados en un conjunto de referencia de objetos de contenido. El control del vocabulario es aplicado desde el principio como una operación continua. Si el vocabulario controlado tiene algún tipo de jerarquía, cada término es asignado a las clases que le correspondan apenas se decide su inclusión. La ayuda de expertos del área es recomendada fuertemente; estos expertos pueden formar parte de un consejo de redacción o comité.
- *Combinación de métodos* - En la práctica, más de un enfoque puede ser empleado en una etapa u otra de la construcción de un vocabulario controlado. Por ejemplo, las jerarquías y otras relaciones entre términos que fueron establecidos de forma inductiva pueden ser examinadas luego desde un punto de vista deductivo.

2.7 ANÁLISIS DE FACETAS

Los vocabularios controlados (especialmente aquellos con miles de términos) son mejores de usar si están organizados de otra forma diferente a la jerárquica. El análisis de facetas [Ranganathan, 1933, 1967] es una manera alternativa de organizar el conocimiento. Utiliza un enfoque *bottom-up*, para formar áreas de conocimiento luego de reunir partes del conocimiento y determinar qué áreas forman, en lugar del enfoque *top-down* riguroso de las jerarquías. De esta forma,

se permite asignar múltiples términos para clasificar un objeto de contenido. Por ejemplo, para clasificar una planta se podrían utilizar las siguientes facetas: nombre científico (*pinus pinea*, *juniperus virginiana*, *rosa canina*, etc.), país de procedencia (Uruguay, Estados Unidos, etc.), uso (jardín, medicinal, comestible, etc.). En la práctica, este tipo de análisis involucra identificar los atributos más característicos del contenido a clasificar y seleccionarlos como facetas. El análisis de facetas es particularmente útil para:

1. Nuevas áreas de conocimiento en las cuales el conocimiento del dominio es incompleto o las relaciones entre los conceptos son desconocidas.
2. Áreas interdisciplinarias en las cuales hay más de una perspectiva desde la cual considerar un objeto de contenido.
3. Vocabularios para los cuales se necesitan múltiples jerarquías, pero puede ser inadecuado por la dificultad para definir claramente las fronteras.

TRABAJOS RELACIONADOS

Este capítulo presenta trabajos previos relacionados a la construcción y uso de vocabularios controlados sobre educación de ingeniería de software.

3.1 INTRODUCCIÓN

No fue posible encontrar trabajos previos sobre la construcción de vocabularios controlados sobre educación de ingeniería de software. La búsqueda fue realizada sobre IEEExplore, Science Direct, Scopus y Google Scholar el 25/02/2015, utilizando la cadena «*(taxonomy OR controlled vocabullary) AND software engineering education*».

La falta de publicaciones sobre la construcción de vocabularios controlados no implica que no se utilicen esquemas para clasificar material sobre educación de ingeniería de software. Estos esquemas de clasificación carecen, en general, del rigor de un vocabulario controlado. Aún así encontramos importante su relevamiento para conocer el estado del arte en la clasificación de contenido sobre el área temática e identificar posibles características de un vocabulario controlado.

Por tanto, se presenta a continuación un relevamiento de esquemas de clasificación utilizados en libros publicados sobre el tema, en los llamados a trabajos (*call for papers*) de las principales revistas científicas del área y en las revisiones de literatura sobre educación de ingeniería de software. Creemos que estas tres fuentes de material de investigación sobre educación de ingeniería de software son las que pueden brindar mejor información sobre el uso de esquemas de clasificación.

Por último, se incluye la revisión de un artículo sobre la construcción de una taxonomía de palabras clave (*keywords*) sobre el área temática de la investigación de educación de ingeniería.

3.2 LIBROS SOBRE EDUCACIÓN DE INGENIERÍA DE SOFTWARE

En la bibliografía existen varios libros sobre educación de ingeniería de software de los cuales la mayoría corresponden a actas de conferencias [Wasserman and Freeman, 1976] [Gibbs and Fairley, 1986] [Ford, 1988] [Gibbs, 1989] [Dei-

mel, 1990] [Tomayko, 1991] [Sledge, 1992] [Diaz-Herrera, 1994] [Ibrahim, 1995] [Barta et al., 1993] [Inverardi and Jazayeri, 2005] [Inverardi and Jazayeri, 2005]. Estos libros en lugar de usar un esquema de clasificación presentan los artículos sin ninguna agrupación o agrupan artículos según algunos de los tópicos utilizados en los llamados a trabajos de las revistas o conferencias, cuyo relevamiento fue estudiado de forma independiente (presentado en la Sección 3.3). Debido a esto no se presenta aquí un mayor análisis.

Es interesante estudiar el siguiente libro.

Ellis, Heidi J. C; Demurjian, Steven A, and Naveda, J. Fernando. *Software en-gineering effective teaching and learning approaches and practices*. Information Science Reference, Hershey, USA, 2009.

Los editores describen al libro como «una agregación de técnicas y experiencias en el aula cosechado en todo el mundo y que han sido probadas con éxito en la educación de ingenieros de software.» No es un libro de texto sino una recopilación de artículos de distintos investigadores del área. No queda claro el procedimiento de revisión, pero sí se incluye una lista de revisores.

Los artículos son agrupados en un conjunto de categorías que se presentan a continuación; es necesario destacar que el libro no establece una descripción o alcance para cada una. Las descripciones que presentamos fueron deducidas según el nombre de las categorías y aspectos que comparten los artículos que agrupan.

- *Innovative teaching methods* - Investigación que involucra métodos de enseñanza innovadores.
- *Curriculum education management* - Investigación relacionada a la creación, implantación y evaluación de nuevos currículos.
- *Education technology* - Investigación que contempla la aplicación de tecnología.
- *Professional practice* - Investigación relacionada a la práctica profesional: ética, acreditación internacional, etc.
- *Project-based software engineering* - Investigación que involucra a la educación de ingeniería de software basada en proyectos.
- *Student learning assesment* - Investigación relacionada a las teorías de aprendizaje y enseñanza.

Aunque las categorías que utilizan los editores resultan interesantes carecen del rigor de un vocabulario controlado y también tienen un sesgo importante que corresponde al interés de los propios editores. No indican en ningún lugar cómo fueron seleccionadas esas categorías, así como tampoco su descripción o su relevancia. Tampoco es posible determinar cómo fue el proceso de asignación de artículos a las categorías, en algunos casos es difícil saber por qué un artículo pertenece a una categoría y no a otra.

3.3 TÓPICOS EN LAS CONFERENCIAS SOBRE EDUCACIÓN DE INGENIERÍA DE SOFTWARE

Un llamado a trabajos (o *call for papers* [CFP]) es la solicitud de una conferencia o una revista científica que busca artículos o trabajos académicos para su publicación o presentación en base a criterios preestablecidos. En general se incluye una lista de tópicos o temas de los cuales se aceptan trabajos. En la investigación de vocabularios controlados o taxonomías sobre un área temática las listas de tópicos incluida en los llamados a trabajos de conferencias del área resultan interesantes.

En el Cuadro 1 se presenta la evolución de los tópicos de las últimas ediciones de *Conference on Software Engineering Education and Training* (CSEET) y el *Software Engineering Education and Training Track* (SEET) de la *International Conference on Software Engineering* (ICSE). La conferencia CSEET y el track de educación del ICSE fueron seleccionadas para este estudio por ser los más relevantes en el área específica de educación de ingeniería de software. Para este relevamiento se tomaron en cuenta solamente las últimas tres ediciones (años 2013 al 2015), se estima que es una muestra suficiente para evaluar posibles esquemas de clasificación estables. Se supuso muy poco sobre los tópicos encontrados y solamente en los casos que varios tópicos fueran iguales o muy similares se agruparon en uno sólo. Es claro que sería deseable una mejor correspondencia entre tópicos, pero esto no es posible pues no se cuenta con descripciones ampliadas para cada uno.

Cuadro 1: Tópicos de los llamados a trabajos de las últimas ediciones de CSEET e ICSE.

Tópico	ICSE 2015 CSEET 2015	ICSE 2014	ICSE 2013	CSEET 2014	CSEET 2013
Software and systema development	✓				
New best practices for SEET	✓	✓			

Continúa en la siguiente página

Cuadro 1 – Continuación de la página anterior

Tópico	ICSE 2015 CSEET 2015	ICSE 2014	ICSE 2013	CSEET 2014	CSEET 2013
Innovative curriculum or course formats	✓	✓			
Blending software engineering and other disciplines, such as music, art and psychology	✓	✓	✓		
Blending software engineering and other engineering disciplines, such as electrical engineering and bioengineering	✓				
Exploring the widespread adoption of mobile devices for SEET	✓	✓			
Multidisciplinary team development and project management	✓		✓		
Domain engineering	✓		✓	✓	✓
Social and cultural issues	✓			✓	✓
Global and distributed software engineering	✓			✓	✓
Cooperation in education between industry and academia	✓			✓	✓
Methodological aspects of software engineering education	✓		✓		
Continuous education to cope with technological change	✓	✓		✓	
Experiences with the ACM/IEEE Curriculum 2013		✓			
Impact of online education movement on SE curricula		✓			
Innovative methods for SEET in online courses		✓			
Learnable programming		✓			
Integration of software engineering research into teaching and training		✓	✓		
Education and training paradigms of the new century, such as cloud computing or global software development		✓	✓		✓

Continúa en la siguiente página

Cuadro 1 – Continuación de la página anterior

Tópico	ICSE 2015 CSEET 2015	ICSE 2014	ICSE 2013	CSEET 2014	CSEET 2013
Ensuring graduated students meet new industry needs through the understanding of development practices for different environments		✓	✓		
Innovative use of social media for knowledge management		✓			✓
Different experiences and needs in developing countries		✓			
Teaching outside the normal classroom or laboratory setting, such as education collaborations with industry, community groups and public service organizations enhancing student learning			✓		
Educating the educators, introducing innovative methods which can be used to keep the educators up to date			✓		
Socially relevant assignments that demonstrate how software engineering contributes to society. Today's students want to have a feeling that their work is practical and, especially, to help society and environment			✓		
Software engineering competitions, whether these are related to course-based assignments or to competitions run by external organizations such as industry			✓		
Reading skills				✓	
Assessment				✓	
The web as object or as tool				✓	
Communication with clients, peers, etc.				✓	
Open source in education				✓	
<i>Continúa en la siguiente página</i>					

Cuadro 1 – Continuación de la página anterior

Tópico	ICSE 2015 CSEET 2015	ICSE 2014	ICSE 2013	CSEET 2014	CSEET 2013
Modern development methods, and particularly agile or lean methods					✓
Tools and environment, including open source and commercial products					✓
The internet of things					✓
User experience and its impact on development methods					✓
Application tuning for performance, including power consumption, scalability and other key properties					✓
Entrepreneurship and software business issues					✓
Technology in support of education and training					✓

La lista de tópicos es bastante diferente entre las ediciones y el nivel de detalle en las descripciones también, algo que dificulta mayores comparaciones. En algunos casos los términos utilizados permiten reconocer un concepto de la ingeniería de software (por ejemplo, *'domain engineering'*) pero en otros la descripción es bastante confusa. También se detectan casos en los que dos tópicos de una misma edición son demasiado similares como para suponer que van a clasificar efectivamente la investigación del área (por ejemplo, *'impact of online education movement on SE curricula'* y *'innovative methods for SEET in online courses'*). Parece ser que las listas de tópicos están muy influenciadas por los temas de moda que cambian de un par de años a otro. No se detecta la presencia de un esquema de clasificación de tópicos estable ni carente de ambigüedad en los llamados a trabajos de ambos eventos.

3.4 REVISIÓN TERCIARIA SOBRE EDUCACIÓN DE INGENIERÍA DE SOFTWARE

Un estudio de mapeo es un tipo de revisión de literatura que tiene como propósito encontrar y clasificar estudios primarios (o sea, artículos) con respecto a categorías definidas [Kitchenham et al., 2010]. Es posible detectar el uso de vocabularios controlados en las categorías utilizadas en los estudios de mapeo sobre el tema. Con este enfoque se llevó a cabo una revisión terciaria (un rele-

vamiento de estudios de mapeo) sobre educación de ingeniería de software. La búsqueda se realizó el 25/02/2015 sobre los buscadores de IEEExplore, Science Direct, Scopus y Google Scholar. La cadena de búsqueda utilizada fue la siguiente «('software engineering education' AND ('systematic literature review' OR 'mapping study'))» y los estudios encontrados se muestran en el Cuadro 2.

Cuadro 2: Revisión terciaria sobre educación de ingeniería de software

Artículo	IEEExplore	ScienceDirect	Scopus	Google Scholar
Brereton, P.; Turner, M.; and Kaur, R. Pair programming as a teaching tool: a student review of empirical studies. In <i>22nd Conference on Software Engineering Education and Training</i> , Hyderabad, India, 2009.	✓		✓	
Nascimento, D.M.; Cox, K.; Almeida, T.; Sampaio, W.; Almeida Bittencourt, R.; Souza, R., and Chavez, C. Using Open Source Projects in software engineering education: A systematic mapping study. In <i>43th Frontiers in Education Conference</i> , Oklahoma City, USA, 2013.	✓		✓	✓
Marques, Maira R.; Quispe, Alcides, and Ochoa, Sergio F. A systematic mapping study on practical approaches to teaching software engineering. In <i>44th Frontiers in Education Conference</i> , Madrid, Spain, 2014.	✓			
Jiang, Shu; Zhang, He; Gao, Chao. Process simulation for software engineering education. In <i>8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement</i> , Torino, Italy, 2014.			✓	
Malik, Bushra and Zafar, Saad. A systematic mapping study on software engineering education. <i>International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering</i> , 6(11):3343-3353, 2012.				✓

A continuación se presenta un análisis de los esquemas de clasificación utilizados en los artículos encontrados.

Brereton, P.; Turner, M.; and Kaur, R. Pair programming as a teaching tool: a student review of empirical studies. In *22nd Conference on Software Engineering Education and Training*, Hyderabad, India, 2009.

El trabajo presentado en este artículo busca agregar evidencia sobre la eficacia de la programación de a pares en la enseñanza de introducción a la programación. La revisión fue realizada en el marco de un curso de maestría por

estudiantes supervisados por revisores expertos. El artículo se enfoca más en el proceso de enseñanza del proceso de revisión que en los resultados de la revisión como tal. La búsqueda se realiza de forma manual sobre: ACM Portal, IEEE Xplore, Science Direct y Web of Science; se obtienen 28 artículos relevantes.

Sobre los esquemas de clasificación utilizados únicamente se presenta un conteo de la eficacia (o no eficacia) de la programación de a pares en un conjunto de factores (*exam marks, assignment marks, assignment quality, pass/success rate, retention y confidence/enjoyment/attitude*).

Nascimento, D.M.; Cox, K.; Almeida, T.; Sampaio, W.; Almeida Bittencourt, R.; Souza, R., and Chavez, C. Using Open Source Projects in software engineering education: A systematic mapping study. In *43th Frontiers in Education Conference*, Oklahoma City, USA, 2013.

Este estudio tiene como objetivo resumir la información existente sobre cómo los proyectos *open source* han sido usados para facilitar el aprendizaje de la ingeniería de software. La búsqueda se realiza utilizando: Compendex, IEEE Xplore, ACM, Scopus, Springer y Elsevier. Se obtuvieron 2.204 artículos de los cuales fueron preseleccionados 156 y en una segunda etapa se seleccionaron los 53 artículos relevantes para el estudio. Los artículos estudiados cubren el período de 1998 a 2012.

El esquema de clasificación fue desarrollado utilizando etiquetado (*keywording*) de temas relevantes abordados por las preguntas de investigación y refinado durante la lectura de los artículos. Como resultado, definieron las siguientes facetas, cuyas categorías se reproducen en el Cuadro 3.

- Área de la ingeniería de software (*Software engineering area*): Tópico de la ingeniería de software tratado en el estudio (ingeniería de software en general, requisitos, modelado y análisis, etc.), basado en el SWEBOK [Bourque and Fairley,2014]. Para esta faceta los autores no presentan un detalle de las categorías involucradas aunque, en la presentación de los resultados del estudio, se muestran las siguientes categorías en un gráfico: *Development/Construction, Evolution/Maintenance, Testing, Quality/V&V, Design/Architecture, Modelling/Analysis y Software engineering (General)*. El ítem *Does not apply* fue agregado a algunas facetas para manejar situaciones en las cuales ninguna categorización podía ser aplicada.

- Tipo de investigación (*Research type*): El enfoque de investigación utilizado en los artículos. Las categorías de esta faceta fueron adaptadas de un trabajo previo sobre estudios de mapeo [Petersen et al., 2008].
- Enfoque de aprendizaje (*Learning approach*): El enfoque pedagógico que fue utilizado junto con los proyectos *open source* en los cursos de ingeniería de software.
- Perspectiva de evaluación (*Assessment perspective*): La perspectiva desde la cual los estudiantes son evaluados.
- Tipo de evaluación (*Assessment type*): El instrumento de evaluación utilizado para verificar el aprendizaje de los estudiantes.

Cuadro 3: Categorías del esquema de clasificación de Nascimento et al. [2013]

Category	Description
Research Types	
Experience report	Paper describes the authors' personal experience using a particular approach, explaining what and how something has been done in practice. It usually includes lessons learned.
Case study	An empirical inquiry that investigates a phenomenon within its real-life context. Paper may deal with single or multiple cases, may include quantitative evidence, relies on multiple sources of evidence, and benefits from the prior development of theoretical propositions.
Action research	Interactive inquiry process that balances problem solving actions implemented in a collaborative context with data-driven collaborative analysis or research to understand underlying causes enabling future predictions about personal and organizational change.
Experiment/Quasi-experiment	A collection of research designs that use manipulation and controlled testing to understand causal processes. Generally, one or more variables are manipulated to determine their effect on a dependent variable.
Survey	Encompasses any measurement procedures that involve asking questions of respondents, generally by sampling them from a population. Most common types of surveys use questionnaires or interviews.
Opinion paper	These papers express the personal opinion of somebody whether a certain approach is good or bad, or how things should be done. They do not rely on related work and research methodologies.
Solution proposal	A solution for a problem is proposed, the solution can be either novel or a significant extension of an existing technique. The potential benefits and the applicability of the solution are shown by an example or a good line of argumentation.
Philosophical paper	These papers sketch a new way of looking at existing things by structuring the field in form of a taxonomy or conceptual framework.
Literature review	Critical and in-depth evaluation of previous research. It is a summary and synopsis of a particular area of research

Continua en la siguiente página

Cuadro 3 – Continuación de la página anterior

Category	Description
Learning Approaches	
Does not apply	Paper is not related to an experience where a learning approach is needed.
Not specified	Authors do not state the learning approach used.
Active learning (general)	General term that refers to several models of education that focus the responsibility of learning on learners. Usually students engage in higher-order thinking tasks such as analysis, synthesis, and evaluation.
Case-based learning	Approach where students develop skills in analytical thinking and reflective judgment by reading and discussing complex, real-life scenarios. A case is already organized and synthesized for students.
Game-based learning	Learning that involves students in some sort of competition or achievement in relationship to an educational goal. Attempts to increase student motivation by providing a playful environment.
Peer/Group/Team learning	Educational practices in which students interact with other students to attain educational goals. Such approaches enhance the value of interaction and information sharing among peers.
Problem-/Project-/Inquiry-based learning	A collection of approaches that use projects or problems to drive the learning process. Students learn about a subject through the experience of problem solving, by working in groups with the help of facilitators. Assessment is performance-based and authentic.
Studio-based learning	Approach from professional education, where students undertake a project under the supervision of a master designer. It uses a learning cycle of construction, presentation, critique and response, that is repeated until project completion.
Other	Other approaches different from the previous categories.
Assessment Perspective	
Does not apply	Work is not related to an experience where assessment is necessary.
Not specified	No assessment is mentioned in the paper.
Student perspective	Students assess their learning by either self- or peer evaluation.
Faculty perspective	Students are assessed by faculty or teaching assistants.
Product perspective	Specific criteria are defined to assess students products.
Assessment Type	
Does not apply	Work is not related to an experience where assessment is needed.
None	No assessment instrument is mentioned in the paper.
Exams	Students are assessed by means of written exams.
Reports	Students should write a report for assessment.
Software artifacts	Students are assessed through developed software artifacts.
Passing Tests	Automated tests such as unit tests are used to assess performed work.
Interviews	Interviews are conducted to assess learning.
Seminars	Students are assessed by their performance in seminars.
Portfolio	Students should produce a portfolio as result of their work.
Exercises	Students are assessed by means of exercises.
Surveys	A survey is conducted to assess learning.

Marques, Maira R.; Quispe, Alcides, and Ochoa, Sergio F. A systematic mapping study on practical approaches to teaching software engineering. In *44th Frontiers in Education Conference*, Madrid, Spain, 2014.

Este estudio tiene como objetivo identificar y clasificar los artículos que reportan el uso de experiencias prácticas en la educación de ingeniería de software. La búsqueda se realiza durante octubre de 2013 sobre IEEE Xplore, ACM Digital Library, Web of Knowledge (anteriormente ISI Web of Science), Science Direct (Elsevier), SpringerLink y Wiley International. Se obtienen 7.517 artículos de los cuales quedan 247 en el primer filtrado y por último se obtienen los 173 artículos relevantes; los cuales cubren el período de 1982 a 2013.

Al igual que en el trabajo anterior [Nascimento et al., 2013] el esquema de clasificación fue desarrollado utilizando etiquetado (*keywording*) de temas relevantes y refinado durante la lectura de los artículos. Como resultado, definieron las siguientes facetas, cuyas categorías se reproducen en el Cuadro 4.

- Tipo de investigación (*Research type*): El enfoque de investigación utilizado en los artículos. Las categorías de esta faceta fueron adaptadas de dos trabajos previos [Petersen et al., 2008; Nascimento et al., 2013].
- Enfoque de enseñanza (*Teaching approach*): El enfoque pedagógico de enseñanza utilizado en las experiencias prácticas.
- Modelo de proceso de software (*Software process model*): Indica el tipo de proceso de software utilizado para guiar las experiencias.

Cuadro 4: Categorías del esquema de clasificación de Marques et al. [2014]

Category	Description
Research Type	
Experience report	Author's personal experience describing what and how something was done. This usually includes a lesson-learned section.
Case study	Report of investigation of a real-life phenomenon. It may include quantitative evidence, relies on multiple sources of evidence, and also may report one or several cases.
Action Research	Problem solving actions implemented in a collaborative context with data-driven analysis or research, with the intent to enlighten underlying causes enabling future trend prediction.
Experiment/Quasi-Experiment	Manipulation and controlled testing for causal analysis. Normally, one or more variables are manipulated to determine their effect on the dependent variable.

Continúa en la siguiente página

Cuadro 4 – Continuación de la página anterior

Category	Description
Solution Proposal	A solution is proposed and the results of using the proposed solution are reported. An example or a line of argumentation must show the applicability of the solution.
Teaching approaches	
Case Studies	Students develop skills in analytical thinking by reading and discussing complex-life problems.
Learning by doing	There is no formal reported approach stated in the paper, but report the experiences on using an approach where students take on a project and have to address the clash of software engineering theory and practice.
Game based learning	A game development is proposed as a fun way to teach software engineering and catch the students' attention.
Maintenance	Students have to learn how to deal with software created by other developers, and to participate in its evolution or correction.
Open Source	Students have to learn software engineering by participating actively in an open source community or project.
Problem/Outcome Based Learning (PBL/OBL)	These experiences address practical problems as way to drive the learning process. They can also be focused on getting certain outcomes as a way to emphasize a more functional knowledge than a declarative one. Students learn about a subject through the experience of solving a problem, looking for results and by working in groups.
Simulation	Students work on projects that try to simulate a part of a real life project. For instance they have to play a role, deal with a client, address a problem, and work in an environment similar to industrial scenarios.
Traditional	Involves a lecturer standing in front of students, with the course content divided into a number of topical lectures. Typically, a set of practical assignments is used to help students apply the theoretical concepts to the professional work.
Service Learning	Students develop and use their academic skills to address real life problems within their own environment.
Inverted Classroom	The traditional lecture is placed online for students to utilize during their study time. Their classroom time is used for other activities, such as workshops, interactive work time with the instructor, or demonstrations.
Software Process Models	
Agile	The agile models are product-driven processes that propose a software development approach based on customer collaboration, iterative development and technology adaptable to change. Little or no bureaucracy is involved in these processes.
CMM	CMM (Capability Maturity Model) is a general recommendation that identifies key process areas, and also different levels of work maturity for those areas. These experiences involve planning, performing or validating software process improvement using this recommendation as a guideline.
Mbase	Mbase (Model Based Architecting and Software Engineering) is a particular software process model, which is focused on managing the traceability of the project deliverables.
RUP (based)	Works that report the use of RUP (Rational Unified Process) or a variation of it to guide the development of software projects. Some processes derived from RUP are UP (Unified Process) or UPEDU (Unified Process for Education).
<i>Continúa en la siguiente página</i>	

Cuadro 4 – Continuación de la página anterior

Category	Description
TSP/PSP	Papers that report on the use of TSP (Team Software Process) or PSP (Personal Software Process). These processes are focused on self-measuring and managing the development effort during a project.
Traditional	Works that report the use of waterfall, incremental or any other structured approach for software development. Typically these processes are opposite to agile methods in terms of bureaucracy required to obtain a product.
Various	Papers that report the use of more than one process approach simultaneously. For instance, using a structured model for developing a particular software. We however use an agile approach to address each process stage.
Ad-hoc (stated)	Reports that use a specific process approach, which is described in that work, and that do not match with previous categories.
Not specified	Authors do not state the process approach used to address the work.

Jiang, Shu; Zhang, He; Gao, Chao. Process simulation for software engineering education. In *8th ACM/IEEE International Symposium on Empirical Software Engineering and Measurement*, Torino, Italy, 2014.

Este artículo presenta un estudio de literatura que busca evidencia sobre el uso de simulación de procesos de software en la educación de ingeniería de software. La revisión incluye: una búsqueda manual, otra automática sobre buscadores (IEEE Xplore, ACM Digital Library, Science Direct, Springer Link y Wiley International) y por último una búsqueda sobre las referencias. Se revisan 230 artículos para obtener 16 en la primer búsqueda, 1.097 para obtener 19 y se agregan 7 artículos relevantes por su citado. Los artículos cubren el período de 1992 a 2013.

Los esquemas de clasificación utilizados son los siguientes:

- Tipo de aporte al estudio de los simuladores de procesos de software: contribución al desarrollo o evolución / evaluación empírica o aplicación.
- Lugar de publicación (revista científica o conferencia).
- Nombre o familia del simulador de procesos de software.

Malik, Bushra and Zafar, Saad. A systematic mapping study on software engineering education. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 6(11):3343–3353, 2012.

Este artículo encontrado en la búsqueda en Google Scholar fue publicado por *World Academy of Science, Engineering and Technology (WASET)*, una editorial incluida en *Scholarly Open Access - Critical analysis of scholarly open-access publishing*⁴, una reconocida lista de editoriales cuestionadas (por no cumplir los estándares de referato y control de calidad usuales en la comunidad académica). A pesar de esto, y luego de revisarlo cuidadosamente, hemos decidido incluir el artículo en este relevamiento pues plantea un esquema de clasificación un tanto diferente a los anteriores.

El artículo presenta un estudio de mapeo cuyo objetivo es identificar y clasificar los últimos avances en la mejora de la enseñanza de la ingeniería de software. La búsqueda tiene como objetivo relevar artículos publicados durante 2010 y sea realiza sobre: IEEE Xplore y ACM Digital Library. Se obtienen 1.589 artículos de los cuales 77 se identifican como relevantes, los cuales cubren el período enero a octubre de 2010.

Las clasificación de los artículos se presenta según las siguientes facetas:

- Marco de referencia sobre investigación de educación de ingeniería de software, basado en las agrupaciones presentadas por Ellis et al. [2009]. Malik and Zafar [2012] presentan su descripción de cada categoría, las cuales se reproducen en el Cuadro 5.
- Lugar de publicación (revista científica o conferencia).
- País de publicación y su clasificación según taxonomía de 4 capas según exportaciones.
- Nivel del programa de estudios (*secondary education, undergraduate program, graduate program, undergraduate and graduate programs, certification programs*).
- Área de conocimiento según las Guías para currículas de grado de ingeniería de software (SE2004) [JTFFC,2004]. Aquí se incluyen: *All areas, software management, professional practice, software verification and validation, software verification and validation, software design, software quality, software process, system and application specialties, software modeling and analysis, computing essentials, software evolution y mathematical and engineering fundamentals*.
- Área de conocimiento según la Currícula de referencia para posgrados de ingeniería de software (GSwERC) [iSEEC,2009]. Aquí se incluyen: *All*

4 <https://scholarlyoa.com/publishers/>

areas, software engineering management, testing, software design, ethics and professional conduct, system engineering, requirements engineering, software quality, software construction, software maintenance, configuration management y software engineering process.

- Área de conocimiento según el SWEBOK [Bourque and Fairley,2014]. Aquí se incluyen: *All areas, software engineering management, software design, software engineering tools and methods, software quality, software testing, software engineering process, software requirements, KAs of related disciplines, software construction, software maintenance y software configuration management.*

Cuadro 5: Categorías sobre investigación de educación de ingeniería de software basadas en agrupaciones de Ellis et al. [2009] con descripciones agregadas por Malik and Zafar [2012].

Category	Description
Innovative teaching methods	Research based on the varieties of teaching methods which are innovative in their own way. The goal behind innovation should be to improve student's skill set and to enhance their abstract thinking, awareness and learning.
Curriculum and education management (C&EM)	Research based on how a new or revised curriculum has been developed/implemented according to the current needs of SE field. It also includes research based on how other educational management issues have been addressed or improved.
Educational technology (ET)	Research based on how e-learning technologies are improving the education of SE.
Project-based software engineering (PB-SE)	Research based on how to refine student's learning by giving them the opportunity of practically applying their knowledge in project-based software engineering courses including capstone projects.
Student learning and assessment	Research based on how learning and teaching theories can improve student's learning.

En cuanto a los esquemas de clasificación encontrados en la revisión terciaria de literatura encontramos los siguientes hallazgos:

- En todas las revisiones terciarias consultadas la clasificación de artículos se realiza utilizando un análisis de facetas. No se percibe ningún tipo de consenso en el conjunto de facetas.
- Tres de los cinco esquemas de clasificación cuentan con una descripción de su origen: en dos casos [Nascimento et al., 2013; Marques et al., 2014] se utilizó etiquetado (*keywording*) en los temas relevantes abordados por las preguntas de investigación y refinado durante la lectura de artículos

y en otro [Malik and Zafar, 2012] se utilizaron las categorías presentadas por Ellis et al. [2009] y nombres de áreas temáticas de guías sobre ingeniería de software y currículas [JTFCC,2004; iSEEC,2009; Bourque and Fairley,2014]. El etiquetado (*keywording*) permite contar rápidamente con un esquema de clasificación que cubre los artículos a estudiar [Petersen et al., 2008], pero son posibles dos sesgos: sobreajuste al material utilizado y a los conocimientos de los autores (sobre todo en la definición y descripción de los términos relevados).

- Ninguno de los esquemas vistos cuenta con el rigor de un vocabulario controlado ni ha sido utilizado en más de un estudio terciario.

3.5 VOCABULARIOS CONTROLADOS SOBRE EDUCACIÓN DE INGENIERÍA

Aquí se incluye una discusión del artículo "*Development of a Taxonomy of Keywords for Engineering Education Research*". Este artículo nos fue recomendado en la evaluación del trabajo que nosotros presentamos a la revista *European Journal of Engineering Education*. Es necesario notar que gran parte del trabajo reportado en esta tesis ya había sido realizado al momento del análisis que se presenta a continuación.

Finelli, C. J., Borrego, M., and Rasoulifar, G. (2015). *Development of a Taxonomy of Keywords for Engineering Education Research*. *Journal of Engineering Education*, 104(4), 365-387.

El objetivo de este trabajo es la construcción de una taxonomía de palabras clave (*keywords*) sobre el área temática de la investigación de educación de ingeniería. El alcance de la taxonomía es la investigación realizada en Estados Unidos aunque los autores trataron de que fuera un trabajo inclusivo involucrando a 266 individuos de 30 países. El trabajo comprende tres grandes ítems: la creación de la taxonomía, su validación y la creación de una guía de uso.

Como contexto los autores incluyen una lista de iniciativas que buscan la organización del conocimiento del área de la investigación de educación de ingeniería. Los primeros de estos trabajos categorizan conjuntos de artículos publicados en el *Journal of Engineering Education* [Wankat, 1999, 2004; Whittin and Sheppard, 2004]; más tarde, otros estudios analizaron otras revistas: por su lado Jesiek et al. [2011] analizaron varias revistas y conferencias para extraer 38 categorías de términos y Malmi et al. [2013] categorizaron artículos de la *Société Européenne pour la Formation des Ingénieurs (European Society for Engineerin Educa-*

tion [SEFI]) poniendo foco en el proceso de investigación; por último, Xian and Madhavan [2014] analizaron más de 24,000 artículos de revistas y conferencias de educación de ingeniería usando la taxonomía de tópicos de educación de la base de datos del *Education Resources Information Center (ERIC)* para categorizar redes de colaboración entre académicos.

Para la creación de la taxonomía los autores incorporaron a diversos colaboradores, entre los cuales se destacan expertos de Access Innovations, una empresa de gestión de la información que regularmente crea y mantiene taxonomías. Con su ayuda desarrollaron una taxonomía inicial a partir de tres esquemas iniciales obtenidos de analizar diferentes conjuntos de artículos publicados. Esta taxonomía inicial fue refinada en una serie de 5 *workshops* subsiguientes: la edición 2013 de la *American Society for Engineering Education Annual Conference*, la edición 2013 del *Research in Engineering Education Symposium*, un *workshop* virtual con nueve estudiantes de posgrado, la edición 2013 de la *European Society for Engineering Education Annual Conference* y la edición 2013 de la *Frontiers in Education Annual Conference*. En las distintas instancias los involucrados podían practicar sobre un diverso conjunto de artículos y además se trataban de seguir las consideraciones de la guía Z39.19-2005 provista por ANSI/NISO [NISO,2005] para la mejora de la taxonomía.

La taxonomía satisfizo cuatro criterios de validez y fiabilidad: (1) las palabras clave asignadas a un conjunto de resúmenes de artículos pudieron ser reproducidas por varios investigadores, (2) la taxonomía comprende términos que pudieron ser seleccionados como palabras clave para describir 243 artículos de 3 revistas científicas, (3) las palabras clave para esos 243 artículos estaban distribuidas de manera uniforme a través de las diferentes ramas de la taxonomía, y (4) los autores de 31 artículos de conferencia estuvieron de acuerdo con el 90% de las palabras clave asignadas por los investigadores.

Los autores también elaboraron una guía y una aplicación para iOS para facilitar e incentivar el uso de la taxonomía. También realizaron amplia difusión de los resultados (esta difusión incluye publicaciones en el *Australasian Journal of Engineering Education*, el *European Journal of Engineering Education*, el *IEEE Transactions on Education*, el *International Journal of Engineering Education*, y el *Journal of Engineering Education*). Por último, de acuerdo a las prácticas recomendadas en el área temática de las taxonomías, crearon un comité de revisión⁵.

El resultado comprende una taxonomía de 455 términos organizados en 14 ramas de seis niveles que se reproduce, a causa de su tamaño, junto a la guía de uso en el Anexo A.

⁵ URL <http://taxonomy.engin.umich.edu/people/> - última consulta 25/07/2016

3.6 CONCLUSIONES

No fue posible encontrar trabajos que presenten vocabularios controlados sobre educación de ingeniería de software. Se buscaron también esquemas de clasificación relacionados en libros publicados sobre el tema, en los llamados a trabajos de las principales conferencias del área y en las revisiones de literatura sobre educación de ingeniería de software.

Los esquemas encontrados cuentan con diferentes niveles de rigor y alcance según su utilización. En el libro encontrado sobre educación de ingeniería de software [Ellis et al., 2009] se utiliza una lista muy acotada de categorías para agrupar artículos. No hay descripción para las categorías y en algunos casos no es posible determinar por qué un artículo fue asignado a una categoría y no a otra. Las listas de tópicos de los llamados a trabajos de las conferencias varían mucho en las distintas ediciones, no es posible identificar un esquema de clasificación claro, además tienen bastante ambigüedad en los términos. Los esquemas de clasificación utilizados en las revisiones terciarias son los que cuentan con mayor formalidad en su definición, pero sin llegar al rigor de un vocabulario controlado o a utilizar un mismo esquema en más de una revisión. Todos los esquemas de clasificación incluidos en las revisiones terciarias utilizan un análisis de facetas, para su construcción dos de los trabajos utilizan etiquetado (*keywording*), otro utiliza lista de categorías y tópicos de la literatura y el resto no indican el origen del esquema presentado. Aunque los esquemas encontrados no pueden utilizarse como base para la construcción de un vocabulario controlado (más que nada por la falta de rigor en su construcción) dan una idea sobre posibles usos de los vocabularios controlados sobre educación de ingeniería de software, además presentan un nivel de formalidad y estilo del lenguaje que es utilizado por los usuarios. También puede ser interesante considerar términos de los diferentes esquemas como términos candidatos para un vocabulario controlado del área temática.

Nos fue sugerida la lectura del trabajo de Finelli et al. [2015] sobre la construcción de una taxonomía de palabras clave (*keywords*) para la investigación de educación de ingeniería. Este trabajo presenta dos aspectos interesantes para nuestro trabajo: el método de trabajo y la taxonomía resultante. El método de trabajo presenta algunas características que no son compatibles con el contexto de nuestro trabajo: tiene un enfoque basado en el esfuerzo mayormente de expertos, incorpora consultoría específica en organización del conocimiento y parte de varios trabajos previos sobre clasificación de artículos. En cuanto a la taxonomía resultante, está orientada a la clasificación de investigación de educación de ingeniería en general. Puede ser interesante como trabajo a futu-

ro comparar el resultado de nuestro trabajo con la taxonomía de Finelli et al. [2015]. No descartamos que se pueda extender nuestro trabajo a partir de esa comparación.

MÉTODO DE TRABAJO Y APLICACIÓN

En este capítulo se presenta el proceso utilizado en esta tesis para la construcción de vocabularios controlados y se indica el plan de su aplicación para la construcción de un vocabulario controlado inicial sobre educación de ingeniería de software.

4.1 PROCESO DE CONSTRUCCIÓN DE VOCABULARIOS CONTROLADOS

El proceso que proponemos para la creación de vocabularios controlados contempla un enfoque empírico; para su construcción nos basamos fuertemente en la guía Z39.19-2005 provista por ANSI/NISO [NISO,2005] y las recomendaciones de Hedden [2010]. De ambas fuentes se toman lineamientos y actividades propuestas para armar un proceso que cuenta con etapas definidas y sus salidas previstas. Este proceso puede ser utilizado tanto para la creación de un vocabulario controlado como para la revisión, extensión o actualización de un vocabulario ya existente; y cuenta con dos características relevantes:

- *Enfoque empírico* - Este proceso ha sido construido para trabajar con un enfoque empírico (ver Sección 2.6); en particular fue utilizado para la creación de un vocabulario y luego su ampliación manejando, en ambos casos, un método deductivo. No descartamos pueda ser ajustado para un trabajo inductivo. El trabajar con un enfoque empírico establece que la identificación de conceptos tenga como base el estudio de objetos de contenido o material a clasificar. Dentro de este enfoque el método deductivo propone que se extraigan los términos sin intentar controlar ni determinar relaciones entre ellos hasta recolectar una cantidad suficiente. En una segunda etapa se revisan los términos y se establecen relaciones. El control se establece a partir de entonces.
- *Enfoque de proyectos* - Crear, o incluso revisar y mejorar, un vocabulario controlado comprende un conjunto de diversas actividades que varios autores enmarcan bajo la forma de un proyecto. Este enfoque permite manejar de forma eficaz un amplio conjunto de interesados, que generalmente tienen los vocabularios controlados, así como los riesgos asociados a la construcción de un activo que se pretende sea usado a largo plazo.

A continuación se explica en detalle cada una de las etapas del proceso.

4.1.1 Definición de los objetivos y plan de trabajo

En la Figura 6 se presenta el proceso propuesto; los rectángulos azules son etapas del proceso y los ítems representados por una hoja de papel corresponden a productos de trabajo, los cuales son salida de una etapa y entrada de la siguiente. Por ejemplo, el *Plan de trabajo* es un producto obtenido durante la etapa inicial (llamada *Definición de los objetivos y plan de trabajo*) y es entrada de la etapa de *Recolección de vocabulario*.

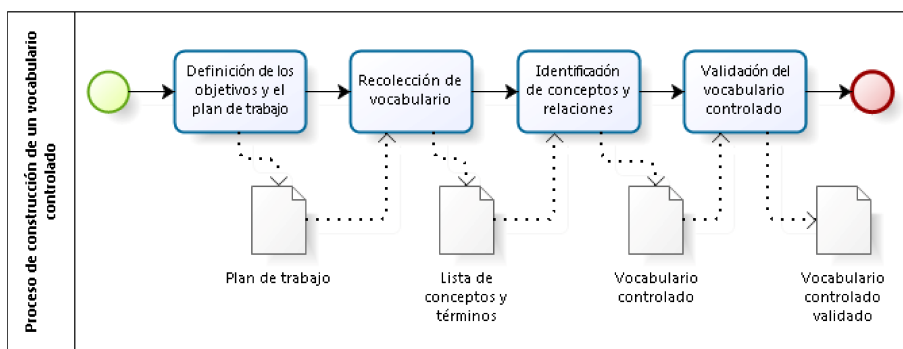


Figura 6: Proceso para la construcción de un vocabulario controlado con un enfoque empírico.

En la primera etapa se definen los objetivos del proyecto, que pueden involucrar la construcción de un vocabulario controlado o la modificación de uno ya existente. En ambos casos se debe incluir el detalle de las características del vocabulario, así como las actividades y recursos para cumplir con los objetivos. Aunque la estructura y el contenido del plan están sujetos al contexto general (por ej, formalidad del proyecto, organización sponsor, etc.) debería incluir las siguientes partes:

OBJETIVO DEL PROYECTO: ¿Cuál es el objetivo del proyecto? Se debe indicar la meta del proyecto a realizar. Ejemplos de objetivos pueden ser: creación o actualización de un vocabulario controlado o la expansión de alguna de sus partes.

PRÓPOSITO DEL VOCABULARIO: ¿Para qué será utilizado el vocabulario controlado? Típicamente existe un propósito principal para un vocabulario

controlado y puede referirse a: soporte para indexación de contenido (humano o automático), soporte para recuperación de contenido o soporte para organización y navegación. El propósito es un factor determinante para establecer la estructura que tendrá el vocabulario, por ejemplo, para un vocabulario cuyo objetivo es la recuperación de información una estructura jerárquica o facetada puede ser ideal. Cuando sea posible se debe indicar que estructura se espera que tenga el vocabulario controlado.

USUARIOS: ¿Quiénes serán los usuarios del vocabulario controlado? Aunque el propósito del vocabulario contempla a quién está dirigido, se deben especificar otras características de los usuarios que puedan influir en el vocabulario a construir. Algunas consideraciones relevantes son: ocupación con respecto al uso del vocabulario (dedicación exclusiva en el uso del vocabulario, experiencia en clasificación de contenido, etc.), especialidad (experiencia en el área temática del vocabulario), rango de edades, ubicación geográfica, etc. También se pueden incluir posibles expectativas de los usuarios sobre la presentación del vocabulario controlado o la interfaz de usuario en caso de que sea informatizado.

CONTENIDO: ¿Qué contenido será cubierto por el vocabulario controlado? Se debe especificar cabalmente el contenido a contemplar por el vocabulario controlado, indicando por ejemplo, tipos de contenido (textos, vídeos, etc), lenguaje del contenido, posibles fuentes para los términos del vocabulario, etc.

ALCANCE: ¿Cuál es el área temática, el alcance y los límites? Es necesario establecer cuáles serán los tópicos cubiertos por el vocabulario controlado. Por ejemplo, si se incluirán nombres de personas o marcas.

RECURSOS: ¿Cuáles son los recursos y restricciones del proyecto? Se deben establecer las habilidades y experiencia de los integrantes del proyecto de creación del vocabulario, el esfuerzo que pueden realizar, así como posibles restricciones de tiempo para el proyecto. También se debe incluir la lista de herramientas a utilizar.

ACTIVIDADES DE INVESTIGACIÓN: ¿Qué actividades de investigación se realizarán para recolectar conceptos candidatos a incluir en el vocabulario controlado?

VALIDACIÓN: ¿Se incluirán actividades de validación del vocabulario controlado? ¿Cuáles? Se entiende por validación elementos que indiquen la adecuación al uso.

Como salida de esta etapa se obtiene un plan de trabajo conforme a los objetivos planteados.

4.1.2 *Recolección de vocabulario*

En esta etapa se reúne todo el vocabulario y conocimiento posible sobre el área temática a cubrir por el vocabulario controlado. Para esto se llevan a cabo una o más actividades de investigación, las cuales tienen como objetivo la recolección de términos candidatos a ser incluidos en el vocabulario; aunque, estas mismas actividades, también se pueden utilizar para dar soporte a la toma de decisiones sobre el tamaño del vocabulario, el alcance, la estructura, la presentación y otras características. Posibles tareas de investigación son:

- Entrevistas en profundidad con los principales interesados, con interesados adicionales y con algunos de los potenciales usuarios del vocabulario controlados.
- Auditoría de contenido. Esto comprende la búsqueda de conceptos significativos en el contenido o material a clasificar. En general, es poco práctico o inviable un análisis de todo el material a clasificar por lo cual es recomendable buscar conceptos en: los títulos, secciones del documento, resúmenes (*abstracts*), tabla de contenidos así como en metadata ya existente. Usualmente, se utiliza una planilla para llevar registro de todos los conceptos candidatos mientras se realiza el análisis del material a revisar.
- Análisis de listas de palabras o frases generadas por herramientas de análisis de texto a partir de contenido de ejemplo.
- Análisis de registros de búsqueda (cuando un motor de búsqueda es utilizado).
- Investigar dentro de las funciones de cualquier método de búsqueda/recuperación de la organización y/o métodos utilizados de búsqueda/recuperación para contenido similar en organizaciones similares.

Como salida de esta etapa se obtiene una lista de potenciales conceptos para incluir en el vocabulario controlado así como información adicional de cada uno.

4.1.3 *Identificación de conceptos y relaciones*

En esta etapa se revisa la lista de conceptos candidatos y se incorporan los que sean adecuados al vocabulario controlado. Evaluar la inclusión de uno o más conceptos puede ser una tarea sumamente compleja. Es necesario balancear distintos aspectos de la calidad del vocabulario a construir; por ejemplo, si se incluyen demasiados conceptos, tratando de lograr mayor cobertura es posible que el vocabulario no sea práctico y usable. Un concepto candidato debería ser un término del vocabulario controlado sí las siguientes preguntas pueden ser contestadas de forma afirmativa:

- ¿El concepto se encuentra dentro del área temática comprendida por el vocabulario controlado?
- ¿El concepto es importante y corresponde a algo que los usuarios buscarían?
- ¿Hay suficiente información sobre el concepto?
- ¿Los usuarios esperan que el concepto sea cubierto?

A menos que se trate de la construcción de anillos de sinónimos, será necesario seleccionar un término preferido para describir cada concepto; los otros términos serán registrados como términos no preferidos. Esto también se realiza en esta etapa utilizando las siguientes recomendaciones como guía:

- Utilizar la redacción que los potenciales usuarios más probablemente vayan a utilizar.
- Mantener foco en el alcance o área temática del vocabulario controlado.
- Imponer un vocabulario controlado organizacional o empresarial.
- Utilizar estándares académicos, culturales, políticos o industriales.
- Mantener consistencia en el estilo a lo largo de todo el vocabulario controlado.
- Utilizar la redacción de los términos comprendidos por el material/contenido a clasificar por el vocabulario controlado.

Las relaciones a identificar y modelar dependen de la estructura y objetivos definidos previamente. A pesar de esto y para evitar la ambigüedad se deben

identificar todas las relaciones de equivalencia (sinónimos) que sea posible. En el caso particular de vocabularios controlados asociados a áreas poco exploradas, como es el caso de educación de ingeniería de software, se recomienda registrar dos términos como relacionados aunque no sea posible determinar la naturaleza de la relación. Esto significa, por ejemplo, que dos conceptos serán registrados como «relacionados» si se tienen dudas si son sinónimos o si uno contiene al otro.

Como salida de esta etapa se obtiene un vocabulario controlado. La estructura y otras características del vocabulario controlado obtenido estarán sujetas a las definiciones realizadas en el plan.

4.1.4 Validación del vocabulario controlado

Validar un vocabulario controlado implica recolectar elementos que indiquen su adecuación al uso. Este proceso no establece tareas o actividades para validación ya que, en general, es difícil validar vocabularios controlados sobre áreas temáticas con propósitos de investigación. En este sentido, se recomienda incluir una reflexión sobre las actividades de validación realizadas (o no realizadas) y su posible relevancia. Como guía para seleccionar actividades de validación se pueden utilizar las siguientes perspectivas: adecuación para etiquetar objetos de contenido (garantía literaria), adecuación para ser utilizado por usuarios (garantía de los usuarios) y adecuación a la organización (garantía organizacional).

4.2 APLICACIÓN

La aplicación del proceso de construcción explicado antes fue realizada en dos instancias.

En la primera instancia se utilizó una herramienta de procesamiento de textos para explorar más de 1.000 artículos y obtener una primera versión de un vocabulario controlado sobre educación de ingeniería de software. Se identificaron y validaron tres facetas como prometedoras para establecer un primer nivel en la jerarquía de términos, estas corresponden a *know-what* ('*what to teach*'), *know-how* ('*how to teach*') y *know-where* ('*where to teach*'). Cada faceta agrupa una lista de términos asociados a los conceptos más investigados a partir del estudio de la frecuencia de términos: '*what to teach*' (24 términos), '*how to teach*' (13 términos) y '*where to teach*' (6 términos).

En la segunda instancia se expandió la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' de forma de incorporar más conceptos rela-

cionados a enfoques y técnicas utilizados para enseñar ingeniería de software. También se utilizaron técnicas de síntesis temática para recolectar y obtener definiciones de todos los conceptos. Se revisaron 34 conceptos candidatos estudiando más de 250 definiciones de distintos autores. El resultado implica la expansión y formalización de la categoría, pasa a incluir 26 términos con sus definiciones y referencias más relevantes. Además se presentan los términos utilizando un formato estándar que permite incorporar toda la información relevante.

En los siguientes dos capítulos se detallan cada una de las dos instancias llevadas a cabo en la construcción del vocabulario controlado sobre educación de ingeniería de software.

VOCABULARIO CONTROLADO INICIAL

Este capítulo presenta la primera instanciación del proceso de construcción de vocabularios controlados descrito en el capítulo anterior para la creación de una taxonomía inicial sobre educación de ingeniería de software.

5.1 MÉTODO

El proceso utilizado es el presentado en el capítulo anterior, su adaptación se muestra en la Figura 7 en la cual se indican las actividades específicas que se realizan en esta instancia para la recolección de vocabulario. El resto de esta sección describe qué aspectos particulares resultan de importancia en cada uno de estos pasos.

5.1.1 *Definición de los objetivos y plan de trabajo*

En esta etapa se definen los objetivos del vocabulario controlado a construir y se establece un plan para su construcción. A continuación se presentan los principales aspectos de nuestro plan de trabajo.

OBJETIVO DEL PROYECTO: El objetivo de esta primera instancia es la construcción de un vocabulario controlado inicial sobre educación de ingeniería de software.

PROPÓSITO DEL VOCABULARIO: El vocabulario controlado tiene como propósito la clasificación de la investigación relacionada a la educación de ingeniería de software. La selección de la estructura depende de las características del vocabulario pero también de los recursos disponibles. Se plantean ciertas dificultades al intentar construir una taxonomía de un área temática poco explorada sin el apoyo de expertos. Las más importantes son sobre la definición de las relaciones jerárquicas, esto es, por ejemplo, cómo definir un concepto como categoría que agrupa a otros. Las relaciones jerárquicas son un factor crítico para que un vocabulario controlado tenga una buena adecuación al uso. Aún con estas dificultades y riesgos se plantea la estructura deseada como taxonomía, pues la

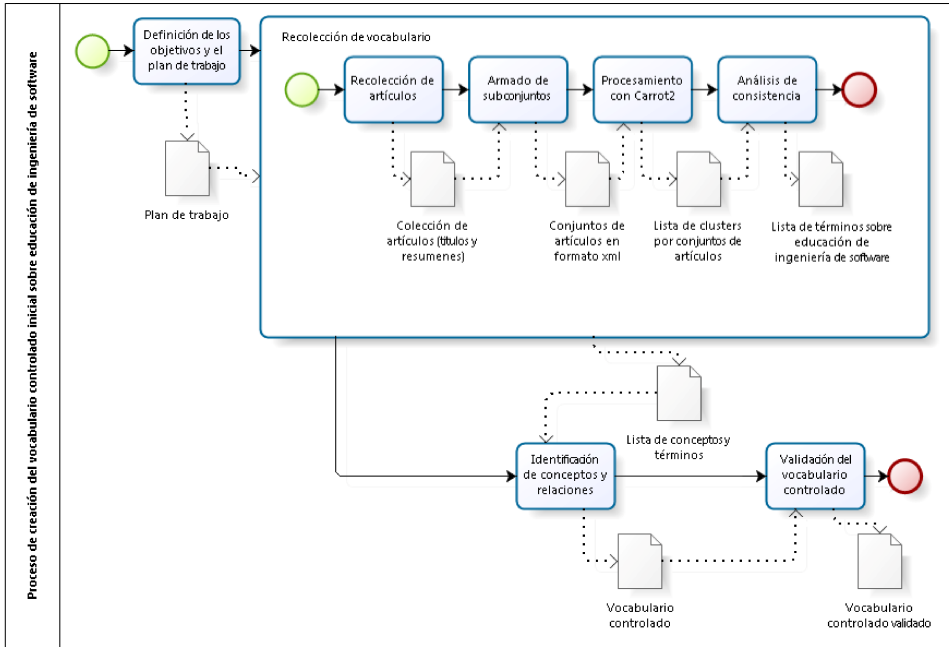


Figura 7: Proceso de creación del vocabulario controlado inicial sobre educación de ingeniería de software.

bibliografía indica que es la estructura más utilizada y con más valor para contextos similares al planteado [Michie et al., 2015; Finelli et al., 2015]. Como se trata de un primer acercamiento al problema, y de forma de minimizar el riesgo de introducir relaciones jerárquicas que degraden la adecuación al uso, se establece como principal objetivo el relevamiento de conceptos y en segundo lugar la definición de relaciones, de este modo se modelarán únicamente las relaciones jerárquicas triviales o aquellas que sean ampliamente respaldadas por la garantía literaria. Para la identificación de las relaciones jerárquicas se priorizará el análisis de facetas.

USUARIOS: Se espera que el vocabulario controlado sea de interés para la comunidad de educación de ingeniería de software. Esto incluye: investigadores, estudiantes, docentes y los editores de revistas que clasifican la investigación. Esta comunidad es dispersa geográficamente y el idioma que más utiliza en la difusión de la investigación es el inglés (en particular, el inglés norteamericano).

CONTENIDO: Debe permitir clasificar el contenido de investigación ya existente (en general artículos o secciones de libros), y su estructura debe soportar su evolución para acompañar la investigación que se realice a futuro.

ALCANCE: Inicialmente el vocabulario será utilizado para clasificar artículos científicos aunque no se desestima su uso en otros tipos de objetos de contenido de investigación. El idioma del vocabulario controlado será el inglés por ser el idioma más utilizado para la difusión de la investigación sobre educación de ingeniería de software.

RECURSOS: El trabajo será realizado por una única persona y no se establecen restricciones de tiempo.

ACTIVIDADES DE INVESTIGACIÓN: Siendo un área poco explorada formalmente en términos de taxonomías y vocabularios controlados se decide priorizar la garantía literaria, con la intención de que el vocabulario cuente con una buena adecuación al uso. La garantía literaria implica que los conceptos incluidos sean los que aparecen realmente en las investigaciones publicadas y que los términos utilizados para representar esos conceptos sean los más utilizados en la literatura. Entonces, como actividad de investigación se plantea la exploración y análisis de material de investigación ya publicado.

VALIDACIÓN: Como se trata de una instancia de exploración que tendrá como resultado un vocabulario controlado inicial y, seguramente, poco completo para un uso práctico no se planifican actividades de validación formales (como puede ser una revisión sistemática de literatura utilizando la taxonomía). A modo de paso tentativo en la validación, se decide planificar la escritura de un artículo que reporte esta primera instancia y sus resultados para la publicación en una revista científica del área. Aunque es claro notar que su publicación no asegura siquiera el posible interés de la comunidad científica en el vocabulario controlado, permite dar visibilidad a la investigación y propicia posibles intercambios.

5.1.2 *Recolección de vocabulario*

La recolección de vocabulario se realiza mediante la técnica de auditoría de contenido (4.1.2), esto es, la búsqueda de conceptos significativos dentro del contenido del material a clasificar. La aplicación de esta técnica para la recolec-

ción de vocabulario involucra 4 pasos que se muestran en el subproceso de la Figura 7.

Como se pretende que los términos del vocabulario sean buenos clasificadores del contenido ya publicado; se elige tener en cuenta el material publicado en la *Conference on Software Engineering Education and Training (CSEET)* durante el período 1988-2014 y en el *Software Engineering Education and Training Track (SEET)* de la *International Conference on Software Engineering (ICSE)* durante el período 2000-2014. Se seleccionan estas conferencias por ser, a nuestro entender, las más reconocidas del área y también porque no existen revistas científicas específicas sobre educación de ingeniería de software. Por ser una instancia de exploración inicial y para lograr cierta garantía literaria decidimos utilizar estas conferencias más específicas en lugar de revistas científicas sobre educación de ingeniería en general (de las cuales sí hay varias importantes). Para ambas se incluyen todos los años disponibles al momento de realizar el trabajo presentado en este capítulo, esto es, todos los años de la CSEET y todos los años en los cuales se publicaron artículos bajo el SEET, hasta el año 2014 inclusive para ambas conferencias.

Es claro que la literatura seleccionada como base para este estudio es relativamente limitada. Existen también otras fuentes de conocimiento sobre educación de ingeniería de software, de las cuales dos importantes son los libros de textos y fuentes de internet como foros de mensajes y grupos informales. Entendemos que este trabajo es un paso inicial en la construcción de una taxonomía sobre educación de ingeniería de software cuyo objetivo principal es la clasificación de la investigación del área. Es por esto que focalizamos nuestro esfuerzo en estudiar el material de investigación detallado anteriormente.

Para la extracción de los términos se utiliza la técnica de análisis de clusters automático (o *clustering*) tomando como base los resúmenes de los artículos. En términos generales, los métodos de análisis de clusters intentan separar los documentos en grupos, donde cada grupo representa un tema que es diferente a los temas representados por los otros grupos [Hammouda and Kamel, 2004; Frakes and Baeza-Yates, 1992]. El *clustering* automático es un método de aprendizaje no supervisado (esto significa que no necesita intervención humana) en el cual se obtiene un conjunto de clusters a partir de un conjunto de documentos o textos (llamado corpus). En general, las técnicas de *clustering* automático se basan en el reconocimiento de patrones sintácticos utilizando la frecuencia de términos similares dentro de los textos para descubrir similitudes. Aunque estas técnicas permiten ahorrar enormemente el esfuerzo a realizar tienen ciertas limitaciones que se deben considerar. Por un lado, la herramienta de *clustering* sólo considera propiedades sintácticas del lenguaje utilizado en los

artículos. De esta forma, se obtiene un conjunto de grupos de documentos que pueden tener o no consistencia semántica. Un análisis del material realizado por expertos tendría en consideración, desde el comienzo, propiedades semánticas obteniendo seguramente más conceptos relevantes y priorizando el foco temático del vocabulario controlado. Por otro lado, los resultados del análisis de clustering automático están sesgados por el algoritmo utilizado por la herramienta. A pesar de estas limitaciones creemos que este tipo de trabajo resulta muy interesante para explorar áreas poco investigadas desde el punto de vista de la organización del conocimiento y los vocabularios controlados.

Como herramienta de *clustering* seleccionamos Carrot² [Weiss and Osinski, 2015]. Carrot² es una herramienta de código abierto para procesamiento de clusters. Fue creada para procesar resultados de búsquedas en internet, pero sus características la hacen una buena herramienta para organizar colecciones de documentos pequeños en categorías temáticas [Weiss, 2006].

El corpus a utilizar en el procesamiento tiene como característica la heterogeneidad de las fechas de publicación de los documentos (desde 1988 hasta 2014). Como la herramienta de *clustering* no utiliza sinónimos, no acumula en un mismo cluster artículos que utilizan diferentes términos para referirse a un mismo concepto. Esto puede convertirse en un sesgo importante ya que sabemos que desde 1988 hasta la actualidad han habido cambios, guiados por modas, en el uso del lenguaje en la ingeniería de software. Para minimizar este sesgo se procesan los artículos en subconjuntos tomando como criterio de agrupación la cercanía en la fecha de publicación. Los subconjuntos fueron seleccionados de forma arbitraria aunque se trata de lograr conjuntos de artículos de tamaños y cantidad de años similares, estos son: 1988-1994, 1995-1999, 2000-2002, 2003-2005, 2006-2008, 2009-2011 y 2012-2014.

Luego de obtenida la lista de clusters para cada período se estudia a mano la composición de cada cluster encontrado. Esto implica la lectura de los resúmenes de los artículos de cada cluster para determinar si se trata de un concepto relacionado a la educación de ingeniería de software o no. En caso afirmativo el término se agrega a la lista de términos a incluir en el vocabulario controlado.

5.1.3 Identificación de conceptos y relaciones

Por último, se revisa la lista de términos obtenidos para identificar conceptos y establecer términos preferidos para cada concepto, así como sinónimos. En esta primera instancia no se buscan relaciones jerárquicas, aunque sí se realiza una primera clasificación en las siguientes facetas que han sido tomadas del trabajo de Nie et al. [2007]:

1. *Know-what* ('*What to teach*') - Aquí se agrupan los conceptos relacionados a tópicos, habilidades o conocimientos que son enseñados a los estudiantes. Algunos ejemplos son: '*software design*', '*communication skills*', '*software testing*', '*project management*'.
2. *Know-how* ('*How to teach*') - Esta faceta contempla conceptos relacionados a enfoques, técnicas o métodos de enseñanza. Todos los conceptos que explican cómo enseñar. Algunos ejemplos son: '*problem-based teaching*', '*case studies*'.
3. *Know-where* ('*Where to teach*') - Se agrupan aquí los conceptos relacionados al marco en el cual se enseñan conceptos relacionados a la ingeniería de software. Esto puede incluir cursos, talleres, clases de educación secundaria, etc. Algunos ejemplos son: '*software engineering course*', '*master degree in software engineering*'.

Estas facetas permiten considerar las publicaciones desde perspectivas que ordenan los conceptos sin perjuicio de futuras relaciones jerárquicas que se establezcan entre los términos del vocabulario.

El uso de estas facetas tiene cierta base empírica: los primeros resultados no eran demasiado comparables; en particular, los términos de los primeros clusters analizados correspondían a conceptos de distintos tipos que parecían tener distintas naturalezas que pudieron ser agrupadas casi directamente en las facetas antes presentadas. La elección de estas facetas, que inicialmente fue intuitiva siguiendo criterios de organización de material para la enseñanza y utilizando como base la lectura de trabajos previos sobre análisis de dominio y organización del conocimiento, resultó ser bastante buena en la agrupación de los distintos términos recolectados.

5.1.4 *Validación del vocabulario controlado*

Como actividad de validación del vocabulario enviamos un artículo a la *European Journal of Engineering Education (EJEE)* el cual fue aceptado para su publicación. El artículo reporta: la motivación, el método de trabajo presentado en el capítulo 4 y su primera instancia incluida en este capítulo. La aceptación de nuestro artículo así como la reciente reimpresión del artículo de Finelli et al. [2015] muestran un creciente interés de la EJEE en investigaciones sobre estandarización de terminologías en el área temática de la educación de ingeniería.

Creemos que es conveniente realizar a futuro otras actividades de validación más relacionadas al uso del vocabulario. Algunas de estas actividades podrían

ser: clasificación de artículos de investigación utilizando los términos del vocabulario o el estudio del vocabulario por parte de un conjunto de expertos en el área de la educación de ingeniería de software.

5.2 RESULTADOS

El Cuadro 6 resume las cantidades de publicaciones utilizadas identificando su fuente y año de publicación agrupadas en los subconjuntos de estudio. Si algún año no se indica en la tabla para alguna de las publicaciones corresponde a que no se encontraron artículos publicados; en particular, para la ICSE esto no significa que no se llevó a cabo la conferencia sino que no se publicaron artículos en el track de educación (SEET).

Cuadro 6: Corpus considerado en la construcción del vocabulario controlado.

Conjuntos	Publicaciones	Cantidad de Artículos
2012-2014	CSEET años 2012 (27), 2013 (53) y 2014 (35). ICSE años 2012 (11), 2013 (13) y 2014 (19).	158
2009-2011	CSEET años 2009 (56), 2010 (29) y 2011 (78). ICSE año 2010 (7).	170
2006-2008	CSEET años 2006 (38), 2007 (48) y 2008 (36). ICSE años 2006 (13), 2007 (23) y 2008 (9).	157
2003-2005	CSEET años 2003 (50), 2004 (35) y 2005 (33). ICSE año 2005 (22).	139
2000-2002	CSEET años 2000 (46), 2001 (28) y 2002 (35). ICSE años 2000 (10) y 2001 (6).	125
1995-1999	CSEET años 1997 (21), 1998 (18) y 1999 (20). CSEE años 1995 (43) y 1996 (20).	122
1988-1994	CSEE años 1988 (18), 1989 (19), 1990 (17), 1991 (20), 1992 (34) y 1994 (44).	152
Cantidad Total		1.023

Se procesaron con la herramienta Carrot² un total de 1.023 artículos agrupados en 7 subconjuntos distintos. Se obtuvo para cada subconjunto una lista de clusters que luego fue revisada manualmente. La revisión manual incluyó la lectura de todos los títulos y resúmenes de los artículos dentro de cada grupo (o cluster) para establecer una consistencia semántica; esto es, la existencia de

uno o más conceptos sobre educación de ingeniería de software asociados al grupo. De esa forma se obtuvieron 43 conceptos que pudieron ser agrupados todos en las facetas: *'what to teach'*, *'how to teach'* y *'where to teach'*. Para cada uno de los conceptos se recolectaron los distintos términos que se utilizan en las publicaciones para representarlo, y se designó al más utilizado como término preferido y al resto como sinónimos. La taxonomía inicial con todos sus términos preferidos se muestra en la Figura 8, en el Anexo B se presenta el vocabulario completo (esto es, los términos preferidos pero también sus sinónimos). A continuación se presentan y discuten los resultados obtenidos.

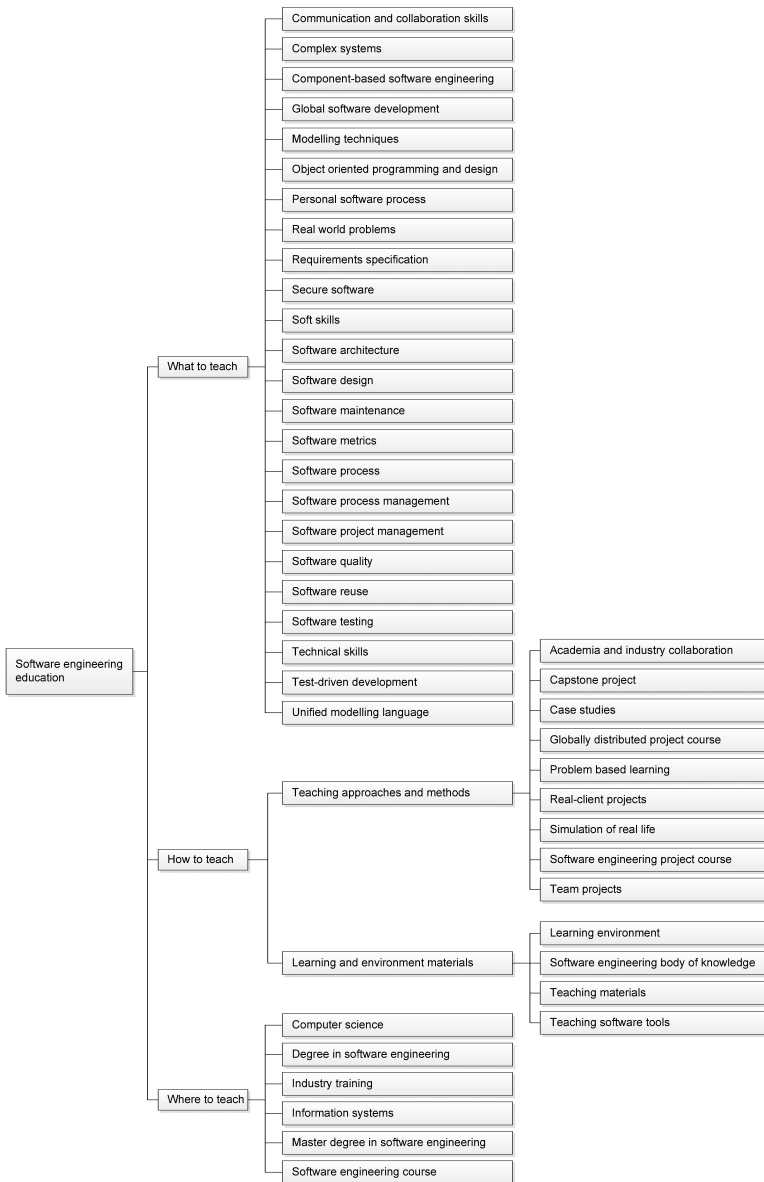


Figura 8: Taxonomía inicial sobre educación de ingeniería de software obtenida en la primera instancia.

5.2.1 *What to teach*

Dentro de la faceta '*what to teach*' se agrupan los términos que representan conceptos relacionados a tópicos, habilidades o conocimientos que son enseñados a los estudiantes. En la Figura 9 se muestra la lista de conceptos de esta faceta (designados por sus términos preferidos) y también se indica para cada concepto los períodos de tiempo en los cuales el análisis de clusters lo detectó como uno de los cuales agrupaba más publicaciones.

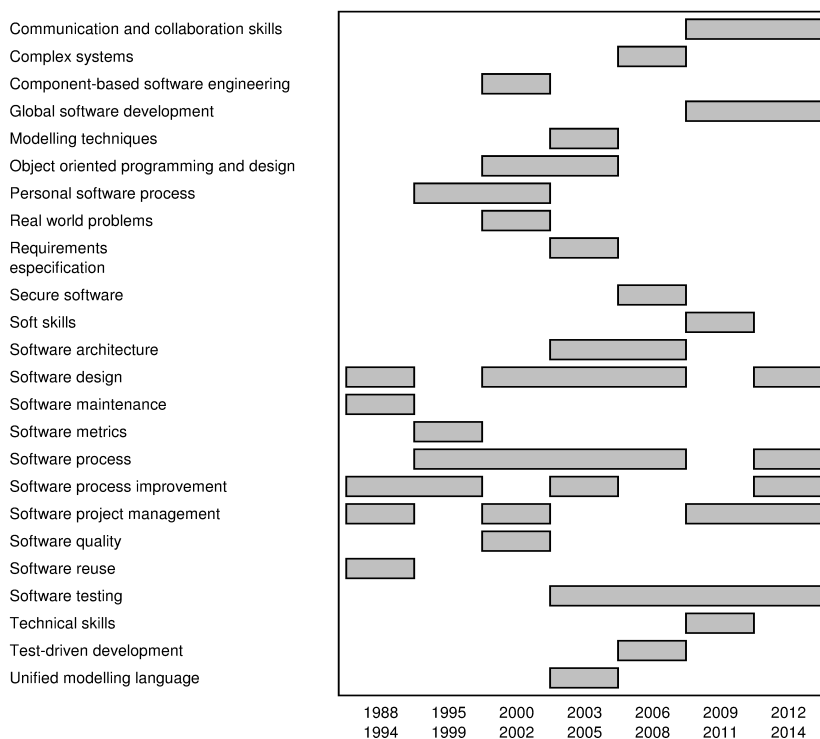


Figura 9: Evolución de los conceptos de la faceta '*what to teach*' a lo largo del tiempo.

Los conceptos asociados a esta faceta contemplan: subdisciplinas de la ingeniería de software ('*software design*', '*software maintenance*', etc.), enfoques de desarrollo ('*object oriented programming and design*', '*test-driven development*', '*personal software process*'), y habilidades ('*communication and collaboration skills*', '*soft skills*', '*technical skills*'). Sobre la evolución, se puede notar cierta diversidad que

aumenta con el tiempo. En los primeros períodos la lista incluía conceptos relacionados a subdisciplinas (en particular *'software design'*, *'software maintenance'*, *'software process'* y *'software project management'*) y enfoques de desarrollo, mientras que en la actualidad también se incluyen conceptos relacionados a habilidades a inculcar. Los conceptos que se han mantenido en más períodos son los siguientes: *'software design'*, *'software process'*, *'software process improvement'*, *'software project management'* y *'software testing'*.

Encontramos interesante estudiar la correspondencia de los conceptos obtenidos con algún material de referencia de ingeniería de software. Seleccionamos para esto la última versión de la guía SWEBOK [Bourque and Fairley, 2014]. *The Guide to the Software Engineering Body of Knowledge (SWEBOK Guide)* es una guía construida por un amplio grupo internacional de profesionales auspiciado por la IEEE-CS⁶, que describe el conocimiento consensuado sobre ingeniería de software. Según sus autores la guía no debe confundirse con el cuerpo de conocimiento en sí mismo, que existe en la literatura publicada. El propósito de la guía es describir la porción del cuerpo de conocimiento que es generalmente aceptado, organizar esa porción, y brindar un acceso por tópicos. El material de la disciplina de ingeniería de software es organizado en una estructura jerárquica, en primer lugar la guía cuenta con 15 áreas de conocimiento, que luego se descomponen en subáreas, tópicos y subtópicos. Para cada tópico o subtópico se incluye una breve descripción y una o más referencias a material que se considera constituye la mejor presentación del conocimiento relativo a ese tópico (o subtópico).

En el marco del estudio de la correspondencia con el vocabulario controlado, cada término de la faceta fue buscado en todo el texto del SWEBOK, las ocurrencias del término (y sus sinónimos) fueron estudiadas para determinar si se trata de un concepto de la guía (esto es: área o subárea de conocimiento, tópico o subtópico o un concepto dentro de un subtópico). Sólo se considera la descripción que presenta la guía de estos conceptos y no se toman en cuenta la lista de referencias asociadas a cada tópico de la guía. En caso de encontrarse un concepto asociado al término de la faceta *'what to teach'* se considera que existe una correspondencia. En caso contrario se considera que no existe correspondencia. Esto mantiene la consistencia con el método utilizado para la construcción de la lista de términos a partir de los clusters de artículos. El resultado del mapeo se incluye en el Cuadro 7.

⁶ El *Institute of Electrical and Electronics Engineers, Computer Society (IEEE-CS)* es una de las sociedades de profesionales de informática más relevantes del mundo.

Cuadro 7: Correspondencia entre los conceptos de la faceta 'what to teach' y el temario del SWEBOK.

Concepto del Vocabulario	Concepto del SWEBOK ⁷
Communication and collaboration skills	11.3 - Communication skills
Complex systems	-
Component-based software engineering	2.7.5 - Component-based design (CBD)
Global software development	-
Modelling techniques	1.4.1 - Conceptual modelling [software requirements] 9 - Software engineering models and methods
Object oriented programming and design	2.7.3 - Object-oriented design 3.4.2 - Object-oriented runtime issues
Personal software process	-
Real world problems	-
Requirements specification	1.5 - Requirements specification
Secure software	13.17 - Secure software development and maintenance
Soft skills	-
Software architecture	2.3 - Software structure and architecture
Software design	2 - Software design
Software maintenance	5 - Software maintenance

Continúa en la siguiente página

⁷ El identificador anterior al concepto indica su ubicación en el temario del SWEBOK, donde el primer número es el capítulo dentro de la guía (que indica el área de conocimiento), el segundo indica subárea y el tercer número indica el tópico.

Cuadro 7 – Continuación de la página anterior

Concepto del Vocabulario	Concepto del SWEBOK
Software metrics	1.7.5 - Measuring requirements 2.5.3 - Measures [software design] 3.2.3 - Construction measurement 4.4 - Test-related measures 5.2.4 - Software maintenance measurement 7.6 - Software engineering measurement 10.3.4 - Software quality measurement 15.3 - Measurement [engineering foundations]
Software process	8 - Software engineering process
Software process improvement	8.3 - Software process assessment and improvement
Software project management	7 - Software engineering management
Software quality	10 - Software quality
Software reuse	3.1.4 - Reuse
Software testing	4 - Software testing
Technical skills	-
Test-driven development	3.4.16 - Test-first programming
Unified modelling language	1.4.1 - Conceptual modelling [software requirements] 3.1.5 - Standards in construction 8.1 - Software process definition 9.2 - Types of models [software engineering models and methods]

Como puede verse la mayoría de los conceptos de la faceta *'what to teach'* encuentran un tópico correspondiente en el SWEBOK. Los conceptos que no tienen correspondencia son: *'global software development'* que podría relacionarse con el tópico 2.7.5 - *Dealing with multicultural environments* pero como el mapeo se basa en la identificación del mismo término o los sinónimos encontrados no podemos considerarlos como correspondientes, conceptos muy generales relacionados a habilidades (*'soft skills'*, *'technical skills'*), conceptos muy abstractos que luego en la evolución de la ingeniería de software fueron cubiertos en temas más concretos (*'complex systems'*, *'real world problems'*) y algún concepto

relacionado a un enfoque de desarrollo específico (*'personal software process'*). En base a esto, no hay nada que indique que el SWEBOK no pueda ser utilizado como referencia para una futura expansión de la faceta *'what to teach'*. Este trabajo de expansión requiere un esfuerzo considerable, ya que es necesario analizar todo el texto del SWEBOK para realizar las actividades de: auditoría de contenido e identificación de conceptos y relaciones (de forma similar a como fueron explicadas en los puntos 5.1.2 y 5.1.3, aunque en este caso el análisis debería ser manual y no siendo conveniente utilizar herramientas de clustering automático). Por esta razón no se presenta como parte de este trabajo sino que se plantea como trabajo a futuro.

5.2.2 How to teach

Dentro de la faceta *'how to teach'* se agrupan los términos que representan conceptos relacionados a técnicas, enfoques o formas de enseñar. En la Figura 10 se muestra la lista de conceptos agrupados en esta faceta y su evolución a lo largo del tiempo.

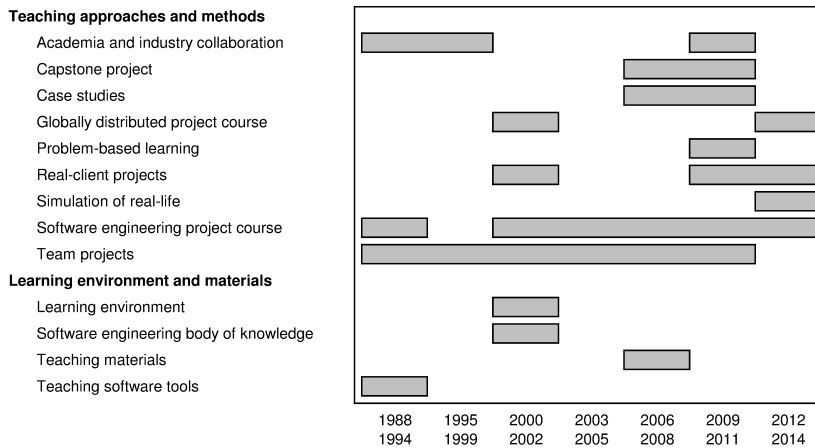


Figura 10: Evolución de los conceptos de la faceta *'how to teach'* a lo largo del tiempo.

Esta faceta incluye gran variedad de tópicos que van desde colaboración entre la industria y la academia hasta materiales de enseñanza. En cuanto a su evolución detectamos que la cantidad de conceptos se expande en los últimos períodos para incorporar más enfoques de enseñanza relacionados a distintas teorías de aprendizaje (*'case studies'*, *'problem-based learning'*, *'simulation of real-life'*). Dos conceptos se mantienen a través de casi todos los períodos, estos son: *'team projects'* y *'software engineering project courses'*. Ambos corresponden a actividades prácticas, de la forma de grupos de estudiantes realizando proyectos (en general de desarrollo). El concepto *'software engineering project course'* involucra un curso específico establecido para esa actividad mientras que *'team projects'* puede involucrar actividades dentro de otro curso (por ejemplo, utilizar *'team projects'* en un curso de *'software design'*).

Es difícil encontrar una base de comparación para la faceta *'how to teach'*. Lo más cercano puede ser el esquema de enfoques de enseñanza presentado por Nascimento et al. [2013], pero no resulta del todo adecuado pues nuestra faceta no refiere únicamente a enfoques de enseñanza. Viendo esto decidimos

clasificar los conceptos de esta faceta en dos subgrupos (los cuales se utilizan para presentar los conceptos de esta faceta en las Figuras 8 y 10):

- *'Teaching approaches and methods'* - En esta categoría se agrupan conceptos relacionados a enfoques y métodos de enseñanza. Algunos ejemplos son: *'case studies'*, *'problem-based learning'*, entre otros.
- *'Learning environment and materials'* - Aquí se agrupan conceptos relacionados a ambiente de aprendizaje y a materiales de enseñanza.

En el Cuadro 8 se presenta la comparación del subgrupo de *'teaching approaches and methods'* y el esquema de clasificación de Nascimento et al. [2013].

Cuadro 8: Correspondencia entre los conceptos de la categoría *'teaching approaches and methods'* y el esquema de clasificación de Nascimento et al. [2013].

Concepto del vocabulario controlado	Categoría del esquema de Nascimento
Academia and industry collaboration	-
Capstone project	-
Case studies	Case-based learning
Globally distributed project course	-
Problem-based learning	Problem-/Project-/Inquiry-based learning
Real-client project	-
Simulation of real-life (environment/problems)	-
Software engineering project course (project courses)	Peer/Group/Team learning
Team projects	
-	Active learning
-	Game-based learning
-	Studio-based learning

Aunque no existe una buena correspondencia, parece adecuado dividir la faceta *'how to teach'* en dos subcategorías: *'teaching approaches and methods'* y *'learning environment and materials'*; y en un futuro ampliar el contenido de ambas para contemplar completamente trabajos como el de Nascimento et al. [2013].

5.2.3 *Where to teach*

Dentro de la faceta *'where to teach'* se agrupan los términos que representan conceptos relacionados al marco en el cual se enseñan tópicos de ingeniería de software. Esto en general puede ser un curso, una carrera académica, un entrenamiento a personal de la industria, etc. En la Figura 11 se muestra la lista de conceptos que incluye esta faceta y su evolución a lo largo del tiempo.

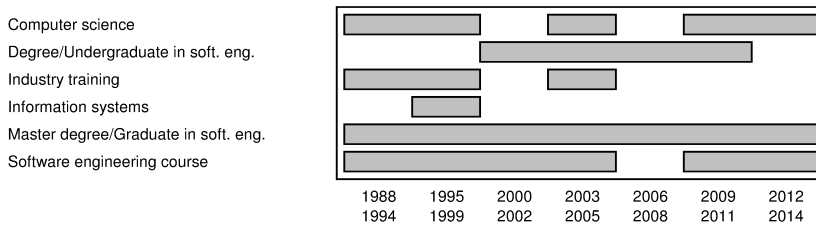


Figura 11: Evolución de los conceptos de la faceta *'where to teach'* a lo largo del tiempo.

En esta faceta se incluyen los siguiente conceptos:

- *'Computer science'* - Enseñanza de ingeniería de software en el ámbito de la ciencia de la computación, esto puede referirse a cursos, departamentos, programas o currículos de esa disciplina.
- *'Degree in software engineering'* - Enseñanza de ingeniería de software en el marco de una carrera de grado en esa disciplina.
- *'Industry training'* - Enseñanza de ingeniería de software en el marco industrial, o sea a trabajadores o profesionales de la industria. Esto en general refiere a capacitación específica o a medida.
- *'Information systems'* - Enseñanza de ingeniería de software en el ámbito de sistemas de información, esto puede referirse a cursos, departamentos, programas o currículos de esa disciplina.
- *'Master degree in software engineering'* - Enseñanza de ingeniería de software en el marco de un nivel de maestría en esa disciplina.
- *'Software engineering course'* - Enseñanza de ingeniería de software o sus subtópicos en un curso de ingeniería de software.

Sobre la evolución de estos conceptos podemos notar que antes se publicaban experiencias de enseñanza de ingeniería de software en carreras de ciencias de la computación, sistemas de información e ingeniería de software mientras que actualmente se reportan más investigaciones en carreras de ingeniería de software (grado, posgrados o cursos). Esto podría deberse a una mayor formalización del área y de la educación de la ingeniería de software. También se puede ver la importancia que tuvo durante los primeros períodos la difusión de investigación sobre la educación de ingeniería de software directamente en

el contexto industrial (*'industry training programs'*) que actualmente no se mantiene en las dos publicaciones estudiadas.

5.3 LIMITACIONES

Los resultados de esta instancia deben ser interpretados teniendo en cuenta las siguientes limitaciones: (a) solamente se estudiaron los títulos y resúmenes de las publicaciones recolectadas, (b) el criterio utilizado para agrupar los artículos para su procesamiento fue arbitrario, a pesar de intentar obtener subconjuntos de cantidad de artículos y cantidad de años de publicaciones similares, (c) el objetivo principal de este trabajo es obtener un vocabulario controlado inicial utilizando los términos más utilizados, con lo cual el resultado no es un vocabulario exhaustivo, (d) publicaciones a partir del 1 de enero de 2015 no fueron tomadas en cuenta, (e) debido a restricciones de recursos humanos las revisiones manuales de los resultados del análisis de clusters fueron realizadas por un único revisor; y (f) el conjunto de la literatura estudiada en este trabajo es acotado, en particular corresponde a una muestra de artículos de investigación, pues el objetivo de la taxonomía creada es la clasificación de la investigación de educación de ingeniería de software. Existen otras fuentes utilizadas para la enseñanza de ingeniería de software, como por ejemplo libros de texto y foros de mensajes en la web las cuales por falta de recursos y prioridad no han sido tenidas en cuenta.

5.4 CONCLUSIONES

Esta instancia presenta la construcción de un vocabulario controlado inicial para dar soporte a la clasificación y recuperación de investigación sobre educación de ingeniería de software, y por ende, da soporte a la investigación en sí misma. Se muestran los resultados de aplicar técnicas de clustering automático a los artículos publicados en la *Conference on Software Engineering Education and Training (CSEET)* durante el período 1988-2014 y en el *Software Engineering Education and Training Track (SEET)* de la *International Conference on Software Engineering (ICSE)* durante el período 2000-2014 (1.023 artículos) para obtener una lista de conceptos más investigados a partir del estudio de la frecuencia de términos en los resúmenes.

El resultado es un vocabulario controlado en forma de taxonomía con 43 términos que corresponden a los conceptos que se identificaron como los más utilizados en las publicaciones de investigación en el período cubierto. Se presenta la clasificación de los conceptos en tres grandes facetas: *'what to teach'* (24

términos), *'how to teach'* (13 términos) y *'where to teach'* (6 términos). Para cada una de las facetas se muestra también la evolución de los conceptos a lo largo de los años de publicación.

La faceta *'what to teach'* agrupa términos que corresponden a conceptos relacionados a tópicos, habilidades o conocimientos que son enseñados a los estudiantes. Los conceptos relevados contemplan: subdisciplinas de la ingeniería de software (*'software design'*, *'software maintenance'*, etc), enfoques de desarrollo (*'object oriented programming and design'*, *'test-driven development'*, *'personal software process'*), y habilidades (*'communication and collaboration skills'*, *'soft skills'*, *'technical skills'*). Detectamos también un aumento en la diversidad de los conceptos en los últimos años, sobre todo en temas relacionados a habilidades a inculcar. Realizamos una comparación exhaustiva y concluimos que no hay nada que indique que el SWEBOK no pueda ser utilizado como referencia para expandir esta faceta. Para los conceptos que no tienen una correspondencia clara con la estructura del SWEBOK quizás, como trabajo a futuro, sea interesante estudiar el material referenciado por la guía. Es posible que sí exista correspondencia en algunos de esos casos.

Dentro de la faceta *'how to teach'* se agrupan términos que corresponden a conceptos relacionados a enfoques, técnicas o métodos de enseñanza. Todos los conceptos que explican cómo enseñar ingeniería de software. En cuanto su evolución detectamos que la cantidad de conceptos se expande en los últimos períodos para incorporar más enfoques de enseñanza relacionados a distintas teorías de aprendizaje. Identificamos dos conceptos que se mantienen a través de casi todos los períodos (*'team projects'* y *'software engineering project courses'*), ambos corresponden a actividades prácticas en las cuales grupos de estudiantes trabajan realizando proyectos (en general de desarrollo). En relación a esta faceta este trabajo incluye una comparación con el esquema de clasificación utilizado en el trabajo de Nascimento et al. [2013], y aunque no se obtuvieron resultados muy buenos en cuanto a cobertura, permitió detectar la existencia de dos subcategorías bien diferenciadas: *'teaching approaches and methods'* y *'learning environment and materials'*. La gran diferencia entre lo obtenido y la literatura previa sobre los conceptos relacionados a la faceta *'how to teach'* parece indicar la necesidad de complementar nuestro trabajo con análisis de expertos, por lo menos en categorías sin bibliografía de referencia.

La faceta *'where to teach'* modela el marco en el cual se reportan investigaciones sobre educación de ingeniería de software. Esto incluye cursos o carreras relacionadas a otras disciplinas (ciencias de la computación o sistemas de información), capacitaciones a trabajadores de la industria, así como cursos, carreras o programas de ingeniería de software. Lo más destacado es un incremento en

las publicaciones referentes a contextos específicos de ingeniería de software ('*degree in software engineering*', '*master degree in software engineering*' y '*software engineering course*').

Otro trabajo a futuro que parece ser muy interesante es el estudio del vínculo entre las distintas facetas. En este sentido, en la práctica notamos una clara relación que se evidencia, por ejemplo, en la existencia de dos planes de estudio diferentes para nivel de pregrado y carreras de posgrado. Por un lado se cuenta con el *Software Engineering 2014 (SE2014)* [JTFCC,2014], la guía para carreras de pregrado en ingeniería de software que tiene como base el *Software Engineering Education Knowledge (SEEK)* un cuerpo de conocimiento específico; por otra parte está el *Graduate Software Engineering 2009 (GSWE2009)* [iSEEC,2009], la guía para carreras de posgrado, basada también en un cuerpo de conocimiento específico llamado *Core Body of Knowledge (CBOK)*. Ambas guías se basan en cuerpos del conocimiento que tienen como base el SWEBOK (con estrecha correspondencia a la faceta '*what to teach*') a distintos contextos de enseñanza (faceta '*where to teach*').

EXPANSIÓN DE LA CATEGORÍA 'TEACHING APPROACHES AND METHODS'

En su versión inicial el vocabulario controlado sobre educación de ingeniería de software tiene una utilidad bastante acotada; esto se debe principalmente a que cuenta con un conjunto acotado de términos. A esto se le suma no contar con una descripción no ambigua para todos ellos. Para la faceta '*how to teach*' estas limitantes tienen un mayor impacto ya que no se cuenta con material de referencia en la literatura, como puede ser el SWEBOK [Bourque and Fairley, 2014] para la faceta '*what to teach*'. En este capítulo se presenta el trabajo realizado con el propósito de una primera expansión de la categoría '*teaching approaches and methods*' (enfoques y métodos de enseñanza) de la faceta '*how to teach*'.

6.1 MÉTODO

Al igual que en la instancia anterior, el proceso utilizado es el presentado en el Capítulo 4. El resto de la sección describe qué aspectos particulares resultan de importancia en cada uno de los pasos de su instanciación.

6.1.1 Definición de los objetivos y plan de trabajo

En esta segunda instancia se define como objetivo la expansión de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*'. Aunque muchas de las características del vocabulario controlado se mantienen de la versión inicial, se incluyen nuevamente a continuación.

OBJETIVO DEL PROYECTO: El objetivo es la expansión de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' del vocabulario controlado sobre educación de ingeniería de software construido en la instancia anterior. Esto comprende ampliar la cantidad de términos del vocabulario controlado contemplados en la categoría e identificar una definición concisa y lo menos ambigua posible para cada uno.

PROPÓSITO DEL VOCABULARIO: El vocabulario controlado tiene como propósito la clasificación de la investigación relacionada a la educación de

ingeniería de software. La estructura seleccionada es una taxonomía con tres facetas principales.

USUARIOS: La comunidad de educación de ingeniería de software, la cual es dispersa geográficamente y utiliza mayormente el idioma inglés para la publicación de artículos.

CONTENIDO: Debe permitir clasificar el contenido de investigación.

ALCANCE: El vocabulario será utilizado para clasificar artículos científicos aunque no se desestima su uso en otros tipos de objetos de contenido de investigación.

RECURSOS: El trabajo será realizado por una única persona y no se establecen restricciones de tiempo.

ACTIVIDADES DE INVESTIGACIÓN: Se decide mantener la alta prioridad dada en la instancia anterior a la garantía literaria. Como ahora se estudian enfoques y métodos de enseñanza que suponemos son utilizados en distintas disciplinas, no solamente en ingeniería de software, se decide ampliar la revisión de literatura a la que refiera a educación en general.

VALIDACIÓN: Al igual que en la instancia anterior, se decide planificar la escritura de un artículo que reporte esta instancia y sus resultados para la publicación en una revista científica del área. De igual forma, su publicación no asegura siquiera el posible interés de la comunidad científica en el vocabulario controlado pero permite dar visibilidad a la investigación y propicia posibles intercambios.

6.1.2 *Recolección de vocabulario*

En esta instancia, al igual que en la instancia anterior, la recolección de vocabulario se realiza mediante la técnica de auditoría de contenido (4.1.2). En este caso en particular ya se cuenta con una lista inicial de términos en la categoría a estudiar, así que se busca validar ese conjunto y ampliarlo, de forma de que comprenda una mayor cantidad de conceptos. Además, se recopila información específica que luego será sintetizada e incorporada a cada entrada del vocabulario controlado (esto es, para cada término: una definición, posibles sinónimos y referencias relevantes). Las actividades que comprende esta etapa se muestran en la Figura 12 (la cual muestra la adaptación del proceso genérico presentado en el Capítulo 4) y se explican en las siguientes subsecciones.

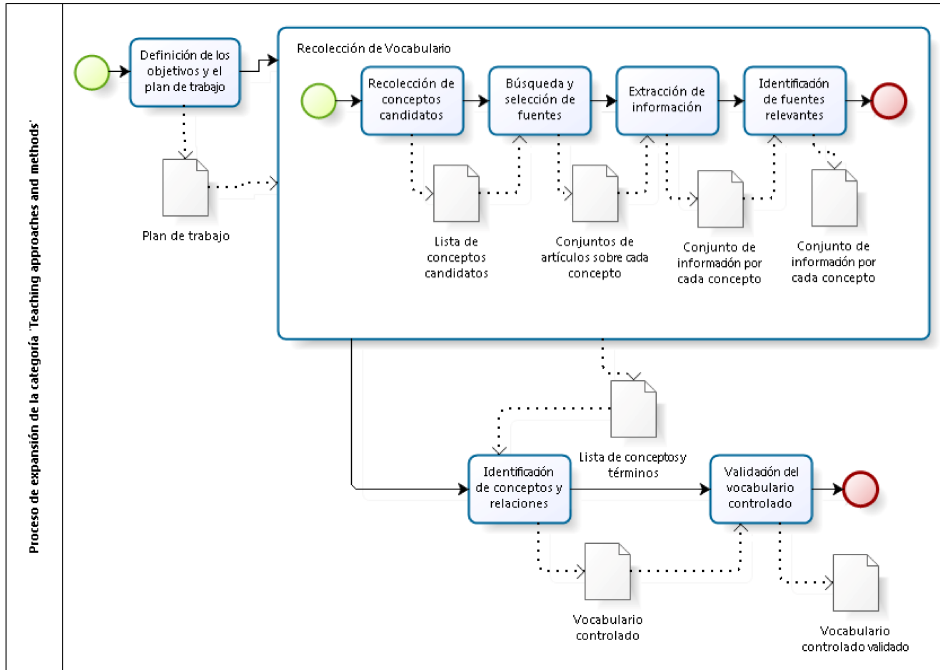


Figura 12: Proceso de expansión de la categoría 'teaching approaches and methods' de la faceta 'how to teach' del vocabulario controlado sobre educación de ingeniería de software.

6.1.2.1 Recolección de conceptos candidatos

En este paso se busca ampliar la lista de conceptos candidatos a incluir en el vocabulario controlado. Para esto se solicita ayuda a la Unidad de Enseñanza de la Facultad de Ingeniería (UEFI)⁸. Con su colaboración se identifican tres fuentes de conceptos relacionados a la enseñanza de ingeniería ([Monte- rrey,2010], [Goodhew, 2014] y [CST,2014]) de las cuales se extraen todos los posibles conceptos candidatos y se agregan a los conceptos obtenidos en la instancia anterior para la categoría 'teaching approaches and methods'. La lista ampliada de conceptos candidatos obtenida se muestra en el Cuadro 9.

⁸ La Unidad de Enseñanza de Facultad de Ingeniería (UEFI) realiza desde su conformación diferentes acciones que buscan estudiar y mejorar los procesos de enseñanza y de aprendizaje en la Facultad de Ingeniería (FIng). La Unidad desarrolla actividades que promueven la formación didáctica de los docentes universitarios así como otras dirigidas a los estudiantes vinculadas con el ingreso y las trayectorias académicas en FIng. Su página web es: <https://www.fing.edu.uy/uefi>

Cuadro 9: Lista ampliada de conceptos candidatos para incluir en la categoría '*teaching approaches and methods*'

Concepto candidato	Tax. inicial	[Monterrey,2010]	[Goodhew, 2014]	[CST,2014]
Lectures		✓		
Tutorials		✓		
Problem-based learning	✓	✓	✓	✓
Project-based learning		✓	✓	
Student-centered learning			✓	
Active learning			✓	
Cooperative learning		✓	✓	
E-learning			✓	
Simulation and games			✓	✓
Distance learning			✓	
Personal development planning			✓	
Case studies	✓	✓		✓
Service-learning		✓		
Research-based learning		✓		
Think-pair-share				✓
Clickers				✓
One minute papers				✓
Interactive lecture demonstrations				✓
Concept mapping				✓
Tutorial worksheets				✓
Just-in-time teaching				✓
Invention activities				✓
Group tests				✓
Problem sets in groups				✓
Random calling				✓
Writing with peer review				✓
Technology enhanced learning			✓	
Academia and industry collaboration	✓			
Capstone project	✓			
Globally distributed project course	✓			
Team projects	✓			
Simulation of real life	✓			
Software engineering project course	✓			
Real-clients projects	✓			

Es interesante notar que de los 9 conceptos obtenidos en la instancia anterior para la categoría '*teaching approaches and methods*' sólo 2 ('*problem-based learning*'

y 'case studies') están contemplados en alguna de las fuentes sobre educación de ingeniería. Esto puede tener diferentes causas. Por un lado, en este paso sólo recolectamos conceptos candidatos en base a los términos que los representan en las distintas fuentes. Esto puede traer cierta redundancia que luego será tratada en los siguientes pasos. Por ejemplo, más adelante se detectó que el concepto 'simulation of real life' se encuentra muy relacionado al concepto 'simulation and games'. Por otro lado, también es posible que existan técnicas y métodos de enseñanza que sean muy específicos de la educación de la ingeniería de software y que no sean tan aplicados a la educación de ingeniería (por ejemplo, el concepto 'software engineering project course' que corresponde a cursos prácticos en los cuales grupos de estudiantes se enfrentan a un proyecto, en general, de desarrollo de software). Con la información disponible en esta etapa no podemos asegurar que exista cierta discrepancia terminológica en la sub-disciplina de la educación de ingeniería de software. Lo que sabemos es que el vocabulario controlado resultado de la instancia anterior es incompleto y muestra los términos con más referencias y también es bastante posible que las otras fuentes de conceptos no sean exhaustivas y listen solamente los métodos o técnicas más utilizadas en la educación de ingeniería en general. Ambas cosas hacen que la intersección de los conjuntos sea de pocos conceptos. Es posible notar que esto también ocurre al comparar otras fuentes entre sí.

6.1.2.2 Búsqueda y selección de fuentes

Este paso tiene como objetivo buscar y seleccionar artículos relacionados a cada concepto candidato que permitan, en los siguientes pasos, validar que el concepto es realmente utilizado en la investigación sobre educación y obtener información para: determinar el término a incluir en el vocabulario controlado, elaborar una definición así como una lista de sinónimos y referencias relevantes.

Para evitar un posible sesgo de especialización que corresponde a la deformación de la técnica por su aplicación a la rama de la educación de ingeniería de software, o a la educación de ingeniería en general, se consideran artículos sobre enseñanza (sin restricciones).

Entonces, para cada concepto listado en el Cuadro 9 se arma una cadena de búsqueda que incluye los siguientes términos: *Teaching approach(es)*, *Teaching method(s)*, *Teaching technique(s)*, *Learning approach(es)*, *Learning method(s)*. En el Cuadro 10 se presenta un ejemplo.

Se realiza luego la búsqueda en SCOPUS y se eligen una cantidad suficiente de artículos de la lista ordenada por cantidad de citas. Luego de algunas pruebas hemos determinado que seleccionar 20 artículos resulta adecuado, se

Cuadro 10: Ejemplo de cadena de búsqueda para el concepto '*Lectures*'.

```
((TITLE-ABS-KEY("lectures") OR TITLE-ABS-KEY("lecture") OR TITLE-ABS-KEY("lecture-based"))) AND
(TITLE-ABS-KEY("teaching approach") OR TITLE-ABS-KEY("teaching approaches") OR TITLE-ABS-
KEY("teaching method") OR TITLE-ABS-KEY("teaching methods") OR TITLE-ABS-KEY("learning ap-
proach") OR TITLE-ABS-KEY("learning approaches") OR TITLE-ABS-KEY("learning method") OR TITLE-
ABS-KEY("learning methods") OR TITLE-ABS-KEY("teaching technique") OR TITLE-ABS-KEY("teaching
techniques") OR TITLE-ABS-KEY("learning technique") OR TITLE-ABS-KEY("learning techniques")))
```

debe recordar que estos artículos serán insumos para los siguientes pasos. El proceso de selección comprende la lectura de los títulos y resúmenes de los artículos tratando de encontrar aquellos que refieran a fundamentos del concepto estudiado, a su aplicación o a su comparación con otras técnicas. Artículos que refieran indirectamente al concepto no son tomados en cuenta.

6.1.2.3 *Extracción de información*

En el paso anterior se obtuvo un conjunto de artículos para cada concepto candidato, este paso tiene como objetivo la extracción de información relevante sobre cada concepto de los artículos recopilados. De cada artículo se lee la introducción, conclusiones y otras secciones que se consideren relevantes para buscar y registrar: descripciones, sinónimos y referencias a publicaciones sobre el concepto estudiado. En parte superior de la Figura 13 se muestra la salida de este paso para el concepto '*Lectures*'. En este caso, y para facilitar la lectura, varios artículos no se muestran pues fueron descartados al no contar con información del concepto investigado.

6.1.2.4 *Identificación de fuentes relevantes*

Este paso tiene como propósito detectar posibles fuentes relevantes para cada uno de los conceptos candidatos y extraer su información para incorporarla al resultado del paso anterior. Consideramos como fuentes relevantes a aquellos artículos, libros, u otras referencias bibliográficas que son citadas frecuentemente en los artículos recolectados en los pasos anteriores para un concepto candidato.

En la práctica esto se resume a las siguientes actividades. Para cada concepto se revisa la lista de referencias registradas en el paso anterior para detectar publicaciones citadas en tres o más artículos. En caso de detectar publicaciones que cumplan esa condición se busca el texto completo y se añade al conjunto de fuentes a analizar para ese concepto, realizando la extracción de información según lo explicado en el párrafo anterior. Siguiendo con el ejemplo, para el

concepto '*Lectures*' se detectó que una referencia relevante era el libro: Bligh, D.A. (1972). *What's the use of lectures?* Harmondsworth, England: Penguin. Este libro fue analizado y la información relevante incluida a la recolectada en el paso anterior. Esto se muestra en la parte inferior de la Figura 13.

Artículo	Descripción	Sinónimos	Referencias
P. Goodhew, «Teaching engineering», UKCME, Liverpool, 2014.	Se nombra pero no se da una descripción del tema.	-	<ul style="list-style-type: none"> Bligh, Donald, A., What's the Use of Lectures? 3rd Edition, Penguin Books 1972 Sutherland, P. and Badger, R., (2004) Lecturers' Perceptions of Lectures; J Further & Higher Education 28, 3, 277-28
J. Biggs y C. Tang, Teaching For Quality Learning At University. McGraw-Hill Education, 2011.	"A lecture is where the subject matter expert tells the students about the major topics that make up the discipline or professional area, and what the latest thinking is on a topic or discipline."	-	<ul style="list-style-type: none"> Bligh, Donald, A., What's the Use of Lectures? 3rd Edition, Penguin Books 1972
Butler J.A. Use of teaching methods within the lecture format	"The continuous exposition by one person for approximately 50 minutes to a largely passive recipient audience."	-	<ul style="list-style-type: none"> Bligh, Donald, A., What's the Use of Lectures? 3rd Edition, Penguin Books 1972 H. HASEKURA, H. FUKUSHIMA, y K. HIRAIDE, «Evaluation of lectures on medical ethics through students' drawings», Medical Education, vol. 24, n.º 1, pp. 42-45, 1990.
Van Dijk L.A., Jochems W.M.G. Changing a traditional lecturing approach into an interactive approach: Effects of interrupting the monologue in lectures	"Typically consisting of a monologue delivered by the lecturer."	-	<ul style="list-style-type: none"> C. C. Bonwell and J. A. Eison, Active Learning: Creating Excitement in the Classroom, The George Washington University, School of Education and Human Development, Washington, DC, ASHE ERIC Higher Education Report No. 1 (1991). A. Saroyan and L. S. Snell, Variations in lecturing styles, Higher Education, 33(1997) pp. 85-104
Omelicheva M.Y., Avdeyeva O. Teaching with lecture or debate? Testing the effectiveness of traditional versus active learning methods of instruction	<p>"The lecture format is conducive to direct reproduction of information laid out by the instructor. Since the lecturer has maximum control over the flow of information, and is able to provide students with precise, up-to-date, and not otherwise available information more effectively than debaters, who may deliberately or inadvertently misrepresent or omit data, we expect that lecture will be more effective than educational debate in transmitting factual knowledge to students.</p> <p>In a lecture class, students are passive learners, engaged in extensive note-taking."</p>	-	<ul style="list-style-type: none"> Beard, R. Research into Teaching Methods in Higher Education, mainly in British Universities. 4th Ed. Guildford Society for Research into Higher Education, ; 1978. Bligh, D. What's the Use of Lectures?. Harmondsworth, Penguin; 1972. Goolkasian, P. Memory for Lectures. Journal of Educational Psychology. 1979;71. Green, E.B. Relative effectiveness of the lecture and individual readings as methods of college teaching. Genetic Psychology Monographs. 1928;4. Guyot, Y. Teacher location and teacher-student relationships. International Review of Applied Psychology. 1970;19:2. Kinch, W. Recognition Memory. Journal of Experimental Psychology. 1977;3. Lloyd, D.H. A concept of improvement of learning responses in the taught lesson. Visual Education. 1968; (Oct 1968). Thomas, E.J. The variation of memory with time of information appearing during a lecture. in: Studies in Adult Education. 1. University of Bristol, ; 1972.
Fuentes Relevantes			
Bligh, Donald, A., What's the Use of Lectures? 3rd Edition, Penguin Books 1972	"[lectures] are in fact more or less continuous expositions by a speaker who wants the audience to learn something."		

Figura 13: Ejemplo de información extraída para el concepto 'Lectures'

6.1.3 Identificación de conceptos y relaciones

Durante esta etapa se analiza la información extraída y se identifican los conceptos y las relaciones que serán incluidas en el vocabulario controlado. Esta actividad incluye revisar el material y, en primer lugar, descartar conceptos que no tienen artículos de investigación asociados. Para los conceptos con información recolectada el objetivo es establecer una definición, y como se cuenta, en general, con varias fuentes es necesario realizar algún proceso de síntesis de información. Para realizar la síntesis de las definiciones utilizamos una adaptación del proceso de síntesis temática [Cruzes and Dybå, 2011; Creswell, 2011] como se muestra en la Figura 14. Con el objetivo de lograr una adaptación lo más ajustada posible toda la información recolectada en el paso anterior para un concepto fue considerada de la misma forma, esto es, no se trató diferente la información extraída de fuentes identificadas como relevantes.

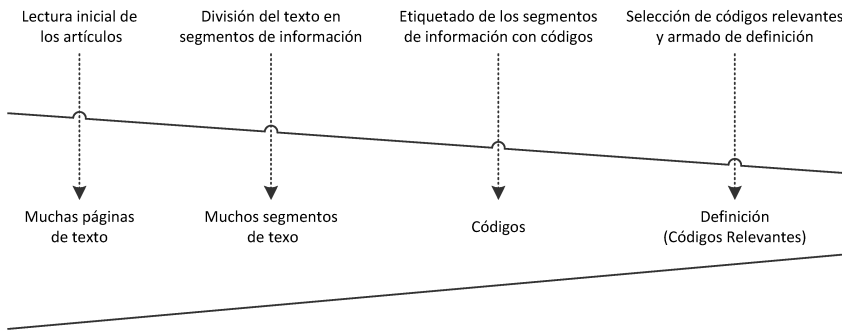


Figura 14: Proceso de síntesis temática adaptado de Creswell [2011]

El proceso corresponde a un tipo de síntesis cualitativa y permite llegar a una definición concisa de cada concepto utilizando los conceptos clave (códigos) que utilizan distintas definiciones de diferentes fuentes de información. En la Figura 15 se presenta la aplicación del proceso de síntesis al concepto *Lectures*. En primer lugar se muestran los segmentos de texto extraídos en el paso anterior de los artículos seleccionados, en la columna de la derecha se muestran los códigos detectados y en la parte inferior se presenta la definición obtenida a partir de los códigos más encontrados. Todas las fases de la síntesis temática, así como todas las actividades presentadas en este capítulo, fueron realizadas de forma manual.

Artículo	Códigos
<p><u>J. Biggs y C. Tang, Teaching For Quality Learning At University. McGraw-Hill Education, 2011.</u></p> <p>A lecture is where the subject matter expert tells the students about the major topics that make up the discipline or professional area, and what the latest thinking is on a topic or discipline.</p>	<p>expert tells the students</p>
<p><u>Bligh, Donald, A., What's the Use of Lectures? 3rd Edition, Penguin Books 1972</u></p> <p>[lectures] are in fact more or less continuous expositions by a speaker who wants the audience to learn something.</p>	<p>continuous exposition speaker audience</p>
<p><u>Butler J.A. Use of teaching methods within the lecture format</u></p> <p>The continuous exposition by one person for approximately 50 minutes to a largely passive recipient audience.</p>	<p>continuous exposition one person passive audience</p>
<p><u>Van Dijk L.A., Jochems W.M.G. Changing a traditional lecturing approach into an interactive approach: Effects of interrupting the monologue in lectures</u></p> <p>Typically consisting of a monologue delivered by the lecturer.</p>	<p>monologue lecturer</p>
<p><u>Omelicheva M.Y., Avdeyeva O. Teaching with lecture or debate? Testing the effectiveness of traditional versus active learning methods of instruction</u></p> <p>The lecture format is conducive to direct reproduction of information laid out by the instructor. Since the lecturer has maximum control over the flow of information, and is able to provide students with precise, up-to-date, and not otherwise available information more effectively than debaters, who may deliberately or inadvertently misrepresent or omit data, we expect that lecture will be more effective than educational debate in transmitting factual knowledge to students.</p> <p>In a lecture class, students are passive learners, engaged in extensive note-taking.</p>	<p>direct reproduction instructor</p> <p>passive learners</p>
Definición obtenida	
Lectures: Lectures are (more o less) continuous expositions of a theme by a speaker to a largely passive recipient audience.	

Figura 15: Ejemplo de códigos en el etiquetado de fragmentos de textos sobre el concepto 'Lectures'

6.1.4 Validación del vocabulario controlado

Como actividad de validación del vocabulario estamos trabajando en un artículo a enviar al *Software Engineering Education and Training Track (SEET)* de la *International Conference on Software Engineering (ICSE)* edición 2017. Este artículo tiene como propósito presentar los resultados de la segunda instancia presentada en este capítulo. Creemos de importancia publicar nuestra investigación en alguna de las dos fuentes utilizadas para la base de nuestro trabajo, siendo el track de educación del ICSE una de ellas.

6.2 RESULTADOS

En esta instancia se revisaron 34 conceptos candidatos y se estudiaron más de 250 definiciones de distintos autores. El resultado comprende la expansión y formalización de la categoría '*teaching approaches and methods*', la cual cubre ahora 26 términos con sus definiciones y referencias más relevantes. Asumimos que un concepto es realmente utilizado en la investigación sobre educación si tiene publicaciones asociadas. Los conceptos candidatos que no se incluyeron finalmente en el vocabulario no tenían ese respaldo bibliográfico o bien eran sinónimos de otros conceptos ya incluidos en el vocabulario.

Se incorpora además un formato para la presentación de cada entrada del vocabulario controlado formal que incluye: término preferido, definición, sinónimos, faceta, términos relacionados y comentarios sobre publicaciones relevantes (en el campo de contenido relacionado, *Related content*). Una entrada de ejemplo se muestra en el Cuadro 11. La taxonomía resultante de esta etapa con todos sus términos preferidos se presenta en la Figura 16 mientras que el detalle de las entradas de la categoría '*teaching approaches and methods*' se muestra en el Anexo C.

Cuadro 11: Entrada de ejemplo con el término *Lectures*

Term	Lectures
Definition note	Lectures are (more or less) continuous expositions of a theme by a speaker to a largely passive recipient audience.
Synonyms	
Facet	How to teach
Related terms	
Related content	The definition included here is based on the classic work: Bligh, D.A. (1972). <i>What's the use of lectures?</i> Harmondsworth, England: Penguin. A new version of this book is available as: Bligh, D. A. (2000) <i>What's the use of Lectures</i> London/San Francisco.

6.3 LIMITACIONES

Los resultados deben ser interpretados teniendo en cuenta las siguientes limitaciones: (a) la lista de conceptos seguramente fue sesgada por las fuentes

iniciales de conceptos candidatos, es bastante probable que existan conceptos relacionados a enfoques y métodos de enseñanza de ingeniería de software que no estén comprendidos aún en el vocabulario controlado presentado, (b) las fuentes iniciales de conceptos candidatos tratan sobre enseñanza de ingeniería en general y no sobre educación de ingeniería de software, (c) el criterio para seleccionar los artículos a considerar en la búsqueda de definiciones e información sobre cada concepto se basa en la lectura de título y resumen únicamente, (d) publicaciones en otro idioma diferente al inglés no fueron tomadas en cuenta durante las búsquedas relacionadas a cada concepto.

6.4 CONCLUSIONES

Esta instancia comprende una primera expansión de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' del vocabulario controlado sobre educación de ingeniería de software construido anteriormente. Los objetivos específicos son ampliar la cantidad de términos de la categoría e identificar una definición concisa y lo menos ambigua posible para cada uno. Para esto fue necesario definir dos métodos: uno para la recolección de vocabulario en detalle y otro para la síntesis de todo el material de un concepto en una única definición. Ambos métodos, aunque basados en investigación previa, son novedosos en su aplicación a la construcción de vocabularios controlados y pueden ser utilizados en trabajos similares.

Luego de aplicar estos dos métodos para recolectar y estudiar más de 250 artículos sobre 34 conceptos candidatos se obtuvieron 26 términos con sus definiciones y referencias más relevantes. Los resultados se incorporan al vocabulario controlado utilizando un formato que incluye para cada entrada del vocabulario controlado formal: término preferido, definición, sinónimos, faceta, términos relacionados y comentarios sobre publicaciones relevantes.

Creemos que es necesario trabajar más en esta categoría para contemplar la variedad de métodos y técnicas que se reportan en las distintas publicaciones de educación de ingeniería de software. La limitación (b) debe ser estudiada en más detalle: como las listas iniciales utilizadas contemplan conceptos de educación de ingeniería de forma genérica puede ser que algunos de los conceptos no apliquen a la educación de ingeniería de software o que existan conceptos de ésta última que no estén cubiertos en la versión actual del vocabulario. Sería conveniente realizar actividades de validación, por ejemplo, utilizando el vocabulario para clasificar artículos de educación de ingeniería de software y ver qué tanto se ajustan los términos al material.

Por otro lado, es bastante posible que distintos investigadores difieran con las definiciones aquí presentadas lo que se puede ver como algo normal ya que no existen puntos de vista únicos. El trabajo realizado es un aporte en la línea de organización del conocimiento y formalización de la terminología sobre educación de ingeniería de software. Como trabajo a futuro creemos que sería muy interesante utilizar análisis de expertos para lograr terminología que surja del consenso, seguramente utilizando los resultados de este y otros trabajos.

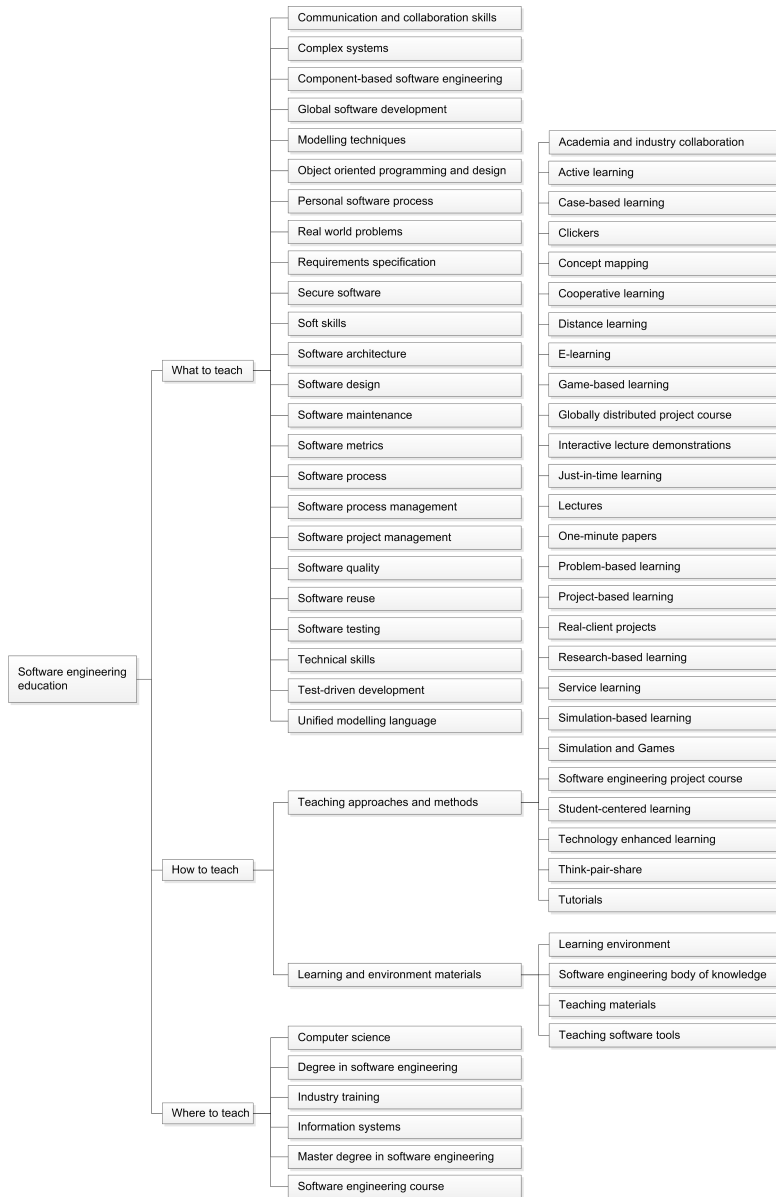


Figura 16: Taxonomía inicial sobre educación de ingeniería de software obtenida en la segunda instancia.

CONCLUSIONES Y TRABAJOS A FUTURO

En este capítulo se presentan las conclusiones de la tesis y se plantean posibles trabajos a futuro.

7.1 CONCLUSIONES

La educación de la ingeniería de software es una disciplina relativamente nueva y es objeto de amplia investigación. Comprende una gran variedad de tópicos y a su vez sufre constantes demandas asociados al rol crítico que tiene el software en el mundo actual. Aunque existen cuerpos de conocimiento y libros de textos sobre ingeniería de software, y más allá de los planes de estudio consensuados por la comunidad, la literatura sobre el área de la educación de ingeniería de software consiste más que nada en artículos publicados en revistas y conferencias especializadas. Este material no se encuentra clasificado en categorías estables ni cuenta con una terminología estándar.

La importancia de organizar el conocimiento y las publicaciones que se realizan está asociada a los beneficios que se obtienen para diferentes involucrados, como son: permitir a los investigadores situar sus iniciativas en un contexto adecuado, lograr una mejor inducción a los términos y conceptos de un área temática, permitir mejorar la gestión de las revistas científicas y conferencias mediante una mejor organización temática de autores e investigación y además mejorar sustancialmente la acumulación de evidencia (revisiones sistemáticas de literatura y estudios de mapeo).

El objetivo general en el cual se enmarca este trabajo es contribuir en la organización del conocimiento sobre el área temática de la educación de la ingeniería de software. En este sentido, el objetivo específico de esta tesis es la creación de un vocabulario controlado sobre educación de ingeniería de software.

En esta tesis se realizó un relevamiento de publicaciones previas sobre la creación y uso de vocabularios controlados y esquemas de clasificación en su definición más amplia relacionados a la educación de ingeniería de software. Aunque no se encontraron trabajos que presenten vocabularios controlados sobre el tema, sí se encontraron esquemas de clasificación en libros publicados sobre educación de ingeniería de software, en los llamados a trabajos de las principales conferencias del área y en las revisiones de literatura. Los esquemas

encontrados cuentan con diferentes niveles de rigor y alcance según su utilización y muestran, a nuestro entender, una mayor necesidad en la formalización del uso del lenguaje. Aunque los resultados del relevamiento no pueden usarse directamente como insumos para la construcción de un vocabulario controlado, principalmente por la falta de rigor en su construcción, permiten deducir algunas características deseables y usos posibles de un vocabulario sobre el tema. Es interesante destacar que todos los esquemas de clasificación son derivados de material a clasificar y son presentados utilizando análisis de facetas.

Partiendo de los lineamientos y actividades propuestas en la guía Z39.19-2005 provista por ANSI/NISO [NISO,2005] y las recomendaciones de Hedden [2010] adaptamos un proceso para la creación y actualización de vocabularios controlados. Este proceso cuenta con etapas bien definidas y salidas previstas y tiene dos características relevantes: contempla un enfoque empírico y su ejecución se basa en proyectos.

Trabajar con un enfoque empírico implica la identificación de conceptos a partir del estudio de objetos de contenido o material a clasificar. Dentro de ese enfoque seleccionamos el método deductivo, el cual propone que se extraigan los términos sin intentar controlar ni determinar relaciones entre ellos hasta recolectar una cantidad suficiente. Luego se revisan los términos y se establecen relaciones, el control se establece a partir de entonces.

Por otro lado, el proceso fomenta el trabajo orientado a proyectos. Un proyecto es un emprendimiento temporal diseñado para producir un único producto con un principio y un fin definidos. Varios autores recomiendan este enfoque de trabajo pues permite manejar de forma adecuada ciertos riesgos asociados a la construcción de un activo que en general debe contemplar un diverso conjunto de interesados y se pretende que sea usado a largo plazo.

El proceso cuenta con 4 etapas bien establecidas: definición de objetivos y plan de trabajo, recolección de vocabulario, identificación de conceptos y relaciones, y validación del vocabulario controlado. Además, establece productos asociados a cada etapa pero no restringe los métodos a utilizar para su ejecución. Esto permite que sea posible su adaptación a diferentes contextos de investigación y recursos disponibles.

Aplicamos el proceso de construcción de vocabularios controlados en dos instancias. Primero para la exploración de literatura existente y creación de un vocabulario inicial sobre educación de ingeniería de software. En la segunda instancia se utilizó el proceso para la expansión de una categoría de una de las facetas identificadas en la instancia anterior. Aunque en ambas instancias se trabajó con el mismo proceso y sobre el mismo vocabulario controlado es importante destacar que los objetivos y las técnicas utilizadas fueron diferentes.

La primera instancia tuvo como objetivo construir un vocabulario controlado para la clasificación de la investigación relacionada a la educación de ingeniería de software. Luego de estudiar los conceptos teóricos asociados a los vocabularios controlados y las características de los esquemas de clasificación utilizados en la literatura sobre educación de ingeniería de software optamos por una estructura taxonómica con análisis de facetas para el vocabulario controlado a construir.

Considerando la suma importancia de la garantía literaria en esta instancia inicial; esto es, la adecuación del vocabulario controlado al material ya existente sobre el área temática, utilizamos como fuentes para la recolección de vocabulario los artículos publicados en las que creemos son las dos publicaciones más reconocidas del área, la *Conference on Software Engineering Education and Training (CSEET)* y en el *Software Engineering Education and Training Track (SEET)* de la *International Conference on Software Engineering (ICSE)*. Para la extracción de los términos utilizamos la técnica de análisis de clusters automático (o *clustering*) y, en particular, la herramienta Carrot² que permite organizar colecciones de documentos pequeños en categorías temáticas. De esta forma fueron procesados los títulos y resúmenes de 1.023 artículos para obtener una lista de términos candidatos a incluir en el vocabulario controlado. Por último fueron analizados de forma manual los grupos encontrados por la herramienta en busca de conceptos y posibles sinónimos.

Como resultado se obtuvo un vocabulario controlado inicial en forma de taxonomía con 43 términos clasificados en tres facetas: *'what to teach'* (24 términos), *'how to teach'* (13 términos) y *'where to teach'* (6 términos). Estas facetas fueron adaptadas del trabajo de Nie et al. [2007] y mostraron ser muy adecuadas como nivel más amplio de la jerarquía.

La faceta *'what to teach'*, que agrupa términos que representan conceptos relacionados a tópicos, habilidades o conocimientos que son enseñados a los estudiantes, fue comparada de forma exhaustiva con los tópicos de la guía SWEBOK [Bourque and Fairley, 2014]. Encontramos que la mayoría de los términos del vocabulario controlado tiene un tópico correspondiente en el SWEBOK, lo que la hacen una muy buena fuente para expandir la faceta *'what to teach'* en un futuro.

La faceta *'how to teach'*, que agrupa términos que representan conceptos relacionados a técnicas, enfoques o formas de enseñar, fue comparada con un trabajo previo de Nascimento et al. [2013] pero no se encontró una correspondencia tan clara. En ese estudio los autores plantean una división en los términos de esta categoría en dos subgrupos: *'teaching approaches and methods'* y *'learning environment and materials'*. Los términos del vocabulario controlado pudieron

ser divididos en esos dos subgrupos fácilmente y creemos que es una buena subclasificación dentro de la taxonomía.

Por último, la faceta *'where to teach'* agrupa términos que representan conceptos referentes al marco en el cual se realiza la educación de ingeniería de software, esto incluye cursos, carreras, entrenamientos a personal de la industria, etc.

Esta instancia inicial entendemos que tiene algunos puntos muy relevantes. El uso de herramientas de procesamiento de lenguaje natural ayuda a explorar un amplio conjunto de objetos de contenido en forma de textos para detectar posibles conceptos a incluir en el vocabulario controlado. Esta exploración permite obtener una lista de términos que cuentan con cierta garantía literaria sin un esfuerzo demasiado excesivo. Además ayuda a validar la primera agrupación de los términos en las tres facetas antes descritas y su evolución a lo largo de los distintos períodos estudiados. Un análisis más minucioso del vocabulario obtenido muestra que la faceta *'what to teach'* puede ser extendida utilizando la guía SWEBOOK mientras que para las otras dos facetas no fue posible establecer bibliografía de referencia.

El método llevado a cabo en esta instancia puede ser utilizado en contextos similares; esto es, para la creación de vocabularios controlados sobre un área temática poco formalizada a partir de exploración de un conjunto grande de objetos de contenido. Se debe considerar que para utilizar este método, y en particular la herramienta Carrot², fue necesario contar con los resúmenes de todos los artículos en formato texto.

Es necesario notar que la literatura seleccionada es limitada y que existen también otras fuentes de conocimiento sobre educación de ingeniería de software, por ejemplo, libros de textos y fuentes de internet como foros de mensajes y grupos informales. A pesar de esto creemos que esta primera instancia cumple con el propósito de lograr una primera exploración de un conjunto representativo del material a clasificar priorizando el foco en la investigación publicada sobre educación de ingeniería de software. Además, existen otras limitaciones relacionadas al método de trabajo: los resultados están sujetos a que solamente se estudiaron los títulos y los resúmenes de los artículos, no es un vocabulario exhaustivo y la revisión fue realizada por un único revisor, entre otras.

La segunda instancia de aplicación del proceso tuvo como objetivo una primera expansión de la categoría *'teaching approaches and methods'* (enfoques y métodos de enseñanza) de la faceta *'how to teach'*. Esta expansión consistió en ampliar la cantidad de conceptos contemplados en la categoría e identificar una definición concisa y lo menos ambigua posible para cada uno.

Se compiló una lista de conceptos candidatos a incluir en la categoría '*teaching approaches and methods*' y luego se aplicaron técnicas de revisión de literatura y de síntesis temática para validar cada concepto y obtener una definición lo menos ambigua posible.

Para realizar la recolección de vocabulario, primero se obtuvo una lista de conceptos candidatos a incluir en el vocabulario controlado; luego, mediante una técnica basada en revisiones de la literatura, se obtuvo para cada concepto distintas definiciones, referencias y sinónimos. Esta técnica permite determinar si un concepto es relevante para un vocabulario controlado a partir de la cantidad y calidad de la literatura encontrada.

Durante la identificación de conceptos fue necesario establecer un mecanismo para obtener una definición lo menos ambigua de cada uno a partir del conjunto de definiciones recolectadas previamente. Para esto se realizó una adaptación del proceso de síntesis temática [Cruzes and Dybå, 2011; Creswell, 2011]. Este proceso que corresponde a un tipo de síntesis cualitativa permite llegar a una definición de cada concepto a partir de un conjunto de definiciones diferentes mediante la extracción de conceptos claves (códigos).

Utilizando ambas técnicas se revisaron 34 conceptos candidatos y se estudiaron más de 250 definiciones de distintos autores. El resultado comprende la expansión de la categoría para cubrir ahora 26 términos con sus definiciones y referencias más relevantes.

El mecanismo de trabajo propuesto permite analizar uno o más conceptos candidatos para su incorporación en un vocabulario controlado ya existente y construir una definición lo menos ambigua posible a partir de la información disponible en bibliografía ya existente. Creemos que aporta el rigor metodológico necesario en la ampliación de un vocabulario controlado manteniendo una base fuerte en la garantía literaria.

Los resultados de aplicar este método se encuentran sesgados por lo incompleta que sea la lista inicial de conceptos candidatos. Otra consideración importante es el alto esfuerzo que lleva realizar la revisión y la síntesis para cada concepto candidato a estudiar.

El vocabulario controlado cuenta con 60 términos organizados en una taxonomía cuyo nivel más alto lo componen tres facetas. A nuestro leal saber este vocabulario es el primero para el área temática de la educación de ingeniería de software y esperamos sea de gran valor para la comunidad. Puede ser utilizado para crear palabras clave a ser utilizadas en el etiquetado de artículos, para entender un concepto y profundizar consultando las fuentes sugeridas y también puede ser usado como base en futuros trabajos de estandarización de la terminología del área. Por la cantidad de conceptos que comprende, cree-

mos que es un vocabulario que debe considerarse como inicial o en proceso de construcción. De ninguna forma debería ser utilizado como si fuera completo o exhaustivo.

El proceso de construcción fue una muy buena guía en ambas instancias. La construcción o actualización de un vocabulario controlado entendemos que tiene ciertas características creativas y de manejo de un gran conjunto de literatura que hacen fácil olvidar la necesidad de rigurosidad a lo largo de todo el proceso. El proceso propuesto es bastante simple pero al establecer etapas y salidas esperadas permite no olvidar ningún paso y lograr la rigurosidad requerida. Puede ser utilizado por otros investigadores para construir vocabularios controlados en distintas áreas temáticas. Creemos que seguir este proceso no tiene un esfuerzo adicional excesivo. La mayor dificultad encontrada, y seguramente le ocurra a otros investigadores, es la identificación y ejecución de actividades de validación que realmente aporten al proceso de construcción realizado. En nuestro caso esto puede tener relación a: que se trata de un vocabulario controlado inicial y que durante esta tesis no contamos con expertos en el área a quienes hacer partícipes.

Los principales aportes de esta tesis son los siguientes:

1. La propuesta de un proceso para la creación de vocabularios controlados con un enfoque empírico y método deductivo y la exploración de dos técnicas, una para la etapa de recolección de vocabulario y otra para la etapa de identificación de conceptos y relaciones.
2. Un vocabulario controlado inicial sobre educación de ingeniería de software que puede ser utilizado por la comunidad para organizar y recuperar información del área. Este vocabulario es una taxonomía de 60 términos clasificados en tres grandes facetas: *'what to teach'*, *'how to teach'* y *'where to teach'*. Para los términos de la faceta *'how to teach'* se incluyen definiciones obtenidas mediante síntesis temática a partir de diverso material consultado.

7.2 TRABAJOS A FUTURO

Posibles líneas de trabajo son:

1. Expandir el vocabulario controlado sobre educación de ingeniería de software presentado.
2. Estudio de las relaciones entre los términos del vocabulario controlado.

3. Utilizar análisis de expertos (o enfoque del comité) para lograr un vocabulario controlado que surja del acuerdo.
4. Validar el vocabulario controlado.
5. Comenzar a utilizar el vocabulario controlado.

El vocabulario controlado obtenido es una aproximación formal inicial a la terminología utilizada en el área de la investigación sobre educación de ingeniería de software. Sería bueno seguir trabajando en su expansión manteniendo el rigor y el respaldo en la garantía literaria. Posibles actividades en esta línea pueden ser: la expansión de la faceta *'what to teach'* utilizando como material de referencia el SWEBOK [Bourque and Fairley, 2014] o continuar la expansión de la categoría *'teaching approaches and methods'* de la faceta *'how to teach'* utilizando, por ejemplo, la reciente taxonomía de *keywords* para educación de ingeniería presentada por Finelli et al. [2015] o explorando métodos y técnicas de enseñanza específicas utilizadas para la educación de ingeniería de software. También puede ser de sumo interés establecer una definición para los conceptos de la taxonomía que aún no cuentan con una.

Hasta ahora el trabajo se ha enfocado en la recolección e identificación de términos y no tanto en el relevamiento y modelado de sus relaciones. En la práctica, hemos identificado solamente tres facetas en el nivel más alto de la jerarquía y luego dos categorías para la faceta *how to teach*. Dentro de esta línea de trabajo parece muy interesante el estudio del vínculo entre las distintas facetas. En este sentido, notamos una clara relación entre la faceta *'what to teach'* a distintos contextos de enseñanza (faceta *'where to teach'*). Esto se puede ver en las guías para currículas de grado y posgrado para la educación de ingeniería de software: el *Software Engineering 2014 (SE2014)* [JTFCC,2014] y el *Graduate Software Engineering 2009 (GSWE2009)* [iSEEc,2009] que tienen como base dos cuerpos de conocimientos (el *Software Engineering Education Knowledge (SEEK)* y el *Core Body of Knowledge (CBOK)*, respectivamente). Ambos diferentes en cuanto a contenido y alcance pero basados en el SWEBOK (con estrecha correspondencia a la faceta *'what to teach'*).

Estudios recientes sobre psicología de la salud presentan un muy buen antecedente en la construcción de una taxonomía, mediante análisis de expertos (o enfoque del comité), cuyo objetivo es clasificar la investigación de un área temática [Michie et al., 2015]. En ese trabajo se parte de varias taxonomías ya existentes. Podemos suponer entonces que el trabajo presentado en esta tesis puede ser visto como un paso preliminar en la construcción de una taxonomía más completa y adecuada al uso utilizando también el enfoque del comité de expertos. En esta línea, creemos que los resultados presentados aquí aportan

cierto orden a la terminología del área temática y que sería beneficioso contar con resultados de análisis similares para luego comenzar un ordenamiento de expertos. Para este trabajo de expertos es posible utilizar como guías los métodos de trabajo planteados por Michie et al. [2015] o Finelli et al. [2015].

Hasta ahora no hemos realizado ninguna actividad de validación del vocabulario controlado (más allá de la publicación del artículo científico sobre la primera instancia de la investigación). Aunque su construcción implica cierta garantía literaria se deben ejecutar actividades que permitan evaluar la adecuación al uso del vocabulario. Es posible, y parece razonable, combinar las actividades de validación con el enfoque de expertos mencionado en el párrafo anterior. Algunas actividades posibles para el vocabulario pueden ser encuestas a expertos o evaluación del uso por posibles usuarios etiquetando artículos científicos. Para lograr avanzar en esta línea es necesario contar con un gran esfuerzo adicional, coordinar trabajos con expertos en el área insume muchos recursos y tiempo. Tenemos como ejemplo el estudio de Michie et al. [2015], el cual implicó un proyecto de 3 años con 400 colaboradores de distintos países para desarrollar y evaluar una taxonomía de 93 técnicas de cambio de comportamiento.

De forma complementaria a la evolución del vocabulario creemos conveniente comenzar con su utilización. En primer lugar estamos trabajando en su difusión, tanto a nivel internacional (mediante los dos artículos de investigación planificados) como en nuestra universidad, mediante un seminario a los docentes del Grupo de Ingeniería de Software (GrIS)⁹. Los docentes del GrIS podrán mejorar el dictado de sus actuales cursos mediante el catálogo de los métodos y técnicas de enseñanza que representa la faceta *'how to teach'*. También dejaremos el vocabulario controlado a disposición de la Unidad de Enseñanza de Facultad de Ingeniería (UEFI) para que pueda ser utilizado por docentes de otras ramas ingenieriles para la mejora de la enseñanza que imparten. Nos interesa también comenzar a utilizar el vocabulario controlado para la clasificación de la investigación sobre educación de ingeniería de software. Siguiendo con la línea de utilizar herramientas de análisis de texto estamos considerando utilizar algoritmos que se puedan entrenar para realizar clasificación automática de grandes conjuntos de artículos utilizando los términos del vocabulario. Esto permitirá contar con investigación clasificada sobre la cual podremos después realizar consultas específicas. Por ejemplo, cruzando dos o tres facetas, para responder a la pregunta ¿Cuáles son los métodos de enseñanza investigados

⁹ El GrIS es el grupo de docentes vinculados a la enseñanza e investigación de ingeniería de software del Instituto de Computación de la Facultad de Ingeniería de la Universidad de la República - <https://www.fing.edu.uy/inco/investigacion/grupos/GRIS>

para el modelado de software (u otros temas)? Este tipo de trabajo de clasificación semi automática (el entrenamiento es con asistencia manual) lleve muchos menos tiempo que la clasificación manual.

BIBLIOGRAFÍA

- Barité, Mario. Garantía literaria y normas para construcción de vocabularios controlados: aspectos epistemológicos y metodológicos. *Scire: Representación y Organización del Conocimiento*, 15(2):13–24, 2009.
- Barité, Mario. *La garantía literaria como herramienta de revisión de sistemas de organización del conocimiento modelo y aplicación*. PhD thesis, Editorial de la Universidad de Granada, Granada, España, 2011.
- Barta, B.Z.; Hung, S.L., and Cox, K.R., editors. *Software engineering education: IFIP WG3.4/SEARCC (SRIG on Education and Training) Working Conference*, International Federation for Information Processing Transactions: Computer Science and Technology, Hong Kong, China, 1993. Elsevier Science.
- Bloom, B.S. *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Taxonomy of Educational Objectives: The Classification of Educational Goals. D. McKay, 1956.
- Bourque, Pierre and Fairley, R. E. *Guide to the Software Engineering Body of Knowledge (SWEBOK)*. IEEE Computer Society, Los Alamitos, USA, 2014.
- Cane, James; Richardson, Michelle; Johnston, Marie; Ladha, Ruhina, and Michie, Susan. From lists of behaviour change techniques (BCTs) to structured hierarchies: Comparison of two methods of developing a hierarchy of BCTs. *British Journal of Health Psychology*, 20(1):130–150, 2015.
- Craig, Edward and Routledge (Firm), , editors. *Routledge encyclopedia of philosophy*. Routledge, London ; New York, 1998.
- Creswell, J.W. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson, 2011.
- Cruzes, D.S. and Dybå, Tore. Recommended Steps for Thematic Synthesis in Software Engineering. In *5th International Symposium on Empirical Software Engineering and Measurement (ESEM)*, pages 275–284, Banff, Canada, 2011.
- CST - Council on Science and Technology, Princeton University. Student-centered teaching methods, 2014. URL <http://www.princeton.edu/cst/teaching-resources/methods/ModifiedPCASTTable.pdf>.

- Deimel, Lionel, editor. *4th Software Engineering Institute (SEI) Conference on Software engineering education*, Pittsburgh, USA, 1990. Springer-Verlag.
- Diaz-Herrera, Jorge L., editor. *7th Software Engineering Institute (SEI) Conference on Software engineering education*, San Antonio, USA, 1994. Springer-Verlag.
- Ellis, Heidi J. C; Demurjian, Steven A, and Naveda, J. Fernando. *Software engineering effective teaching and learning approaches and practices*. Information Science Reference, Hershey, USA, 2009.
- Finelli, Cynthia J.; Borrego, Maura, and Rasoulifar, Golnoosh. Development of a taxonomy of keywords for engineering education research. *Journal of Engineering Education*, 104(4):365–387, 2015.
- Ford, Gary A., editor. *2th Software Engineering Institute (SEI) Conference on Software engineering education*, Fairfax, USA, 1988. Springer-Verlag.
- Foskett, A.C. *The subject approach to information*. London: Library Association, 5th edition, 1996.
- Frakes, William B. and Baeza-Yates, Ricardo, editors. *Information Retrieval: Data Structures and Algorithms*. Prentice-Hall, Inc., Upper Saddle River, USA, 1992.
- Freeman, Peter; Wasserman, Anthony I., and Fairley, Richard E. Essential elements of software engineering education. In *2nd International Conference on Software Engineering*, pages 116–122, Los Alamitos, USA, 1976.
- Gibbs, Norman E., editor. *3th Software Engineering Institute (SEI) Conference on Software engineering education*, Pittsburgh, USA, 1989. Springer-Verlag.
- Gibbs, Norman E and Fairley, Richard E, editors. *Software Engineering Education: The Educational Needs of the Software Community - 1th Software Engineering Institute (SEI) Conference on Software engineering education*, New York, USA, 1986. Springer New York.
- Goodhew, P. Teaching engineering, 2014. URL <http://teachingengineering.liv.ac.uk/book/>.
- Hammouda, K.M. and Kamel, M.S. Efficient phrase-based document indexing for web document clustering. *IEEE Transactions on Knowledge and Data Engineering*, 16(10):1279–1296, 2004.
- Hedden, Heather. *The Accidental Taxonomist*. Information Today, Inc., Medford, USA, May 2010.

- Hjørland, Birger. Domain analysis in information science: Eleven approaches - traditional as well as innovative. *Journal of Documentation*, 58(4):422–462, 2002.
- Hulme, E.W. Principles of Book Classification. *Library Association Records*, 13–14, 1911–1912.
- Ibrahim, Rosalind L., editor. *8th Software Engineering Institute (SEI) Conference on Software engineering education*, New Orleans, USA, 1995. Springer-Verlag.
- Inverardi, P. and Jazayeri, M., editors. *Software Engineering Education in the Modern Age: Software Education and Training Sessions at the International Conference on Software Engineering*, St. Louis, USA, 2005. Springer.
- iSEEc - Integrated Software & Systems Engineering Curriculum project. Graduate Software Engineering 2009 (GSWE2009) Curriculum Guidelines for Graduate Degree Programs in Software Engineering. Technical report, Stevens Institute, 2009.
- Jesiek, B. K.; Borrego, M.; Beddoes, K.; Hurtado, M.; Rajendran, P., and Sangam, D. Mapping global trends in engineering education research, 2005-2008. *International Journal of Engineering Education*, 27:77–90, 2011.
- JTFCC - Joint Task Force on Computing Curricula. Software Engineering 2004: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. Technical report, IEEE Computer Society and Association for Computing Machinery, 2004.
- JTFCC - Joint Task Force on Computing Curricula. Software Engineering 2014: Curriculum Guidelines for Undergraduate Degree Programs in Software Engineering. Technical report, IEEE Computer Society and Association for Computing Machinery, 2014.
- Kitchenham, Barbara A.; Budgen, David, and Brereton, O. Pearl. The Value of Mapping Studies: A Participant-observer Case Study. In *14th International Conference on Evaluation and Assessment in Software Engineering*, pages 25–33, Swinton, UK, 2010.
- Kitchenham, Barbara Ann; Budgen, David, and Brereton, Pearl. *Evidence-Based Software Engineering and Systematic Reviews*. CRC Press, 2015.
- Lambe, P. *Organising Knowledge: Taxonomies, Knowledge and Organisational Effectiveness*. Chandos knowledge management series. Chandos, 2007.

- Lethbridge, Timothy C. What knowledge is important to a software professional? *Computer*, 33(5):44–50, 2000.
- Linstone, Harold A; Turoff, Murray, and others. *The Delphi method: Techniques and applications*, volume 29. Addison-Wesley Reading, MA, 1975.
- Malik, Bushra and Zafar, Saad. A systematic mapping study on software engineering education. *International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering*, 6(11):3343–3353, 2012.
- Malmi, L; Adawi, T; Curmi, R; De Graaff, E; Duffy, G; Kautz, C; Kinnunen, P, and Williams, Bill. Methodological analysis of sefi eer papers. In *41th Conference of the International-Group for the European Society for Engineering Education (SEFI)*, Leuven, Belgium, 2013.
- Marques, Maira R.; Quispe, Alcides, and Ochoa, Sergio F. A systematic mapping study on practical approaches to teaching software engineering. In *44th Frontiers in Education Conference*, pages 1–8, Madrid, Spain, 2014.
- Michie, Susan; Wood, Caroline E; Johnston, Marie; Abraham, Charles; Francis, Jill J, and Hardeman, Wendy. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). *Health Technology Assessment*, 19(99):1–188, 2015.
- Monterrey - Instituto Tecnológico y de Estudios Superiores de Monterrey. Técnicas didácticas, 2010. URL <http://sitios.itesm.mx/va/diie/tecnicasdidacticas/>.
- Nascimento, D.M.; Cox, K.; Almeida, T.; Sampaio, W.; Almeida Bittencourt, R.; Souza, R., and Chavez, C. Using Open Source Projects in software engineering education: A systematic mapping study. In *43th Frontiers in Education Conference*, pages 1837–1843, Oklahoma City, USA, October 2013.
- Nie, Kun; Ma, Tiejun, and Nakamori, Yoshiteru. Building a taxonomy for understanding knowledge management. *Electronic Journal of Knowledge Management*, 5(4):453–466, 2007.
- NISO - ANSI/NISO Z39.19-2005 (R2010). Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies, July 2005 (Reaffirmed May 2010).

- Petersen, Kai; Feldt, Robert; Mujtaba, Shahid, and Mattsson, Michael. Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering*, pages 68–77, Bari, Italy, 2008.
- Ranganathan, Shiyali Ramamrita. Prolegomena to library classification. *The Five Laws of Library Science*, 1967.
- Ranganathan, SR. Colon classification. *Madras, London*, 1933.
- Shaw, Mary. Software engineering education: A roadmap. In *The Future of Software Engineering: 22nd International Conference on Software Engineering*, pages 371–380, Limerick, Ireland, 2000.
- Sledge, C., editor. *6th Software Engineering Institute (SEI) Conference on Software engineering education*, number 640 in Lecture notes on computer science, San Diego, USA, 1992. Springer-Verlag.
- Tennis, Joseph. Subject ontogeny: subject access through time and the dimensionality of classification. *Advances in Knowledge Organization*, 8:54–59, 2002.
- Tomayko, James E., editor. *5th Software Engineering Institute (SEI) Conference on Software engineering education*, Pittsburgh, 1991. Springer-Verlag.
- Wankat, Phillip C. An analysis of the articles in the journal of engineering education. *Journal of Engineering Education*, 88(1):37–42, 1999.
- Wankat, Phillip C. Analysis of the first ten years of the journal of engineering education. *Journal of Engineering Education*, 93(1):13 – 21, 2004.
- Wasserman, Anthony I. and Freeman, Peter, editors. *Interface Workshop on Software Engineering Education: Needs and Objectives*, Irvine, USA, 1976. Springer-Verlag.
- Weiss, Dawid. *Descriptive Clustering as a Method for Exploring Text Collections*. PhD thesis, Poznań University of Technology, Poznań, Poland, 2006.
- Weiss, Dawid and Osinski, Stanislaw. Carrot2 - Open Source Search Results Clustering Engine, 2015. URL <http://project.carrot2.org/>.
- Whittin, Katherine and Sheppard, Sheri. Taking stock: An analysis of the publishing record as represented by the journal of engineering education. *Journal of Engineering Education*, 93(1):5–12, 2004.

Xian, Hanjun and Madhavan, Krishna. Anatomy of scholarly collaboration in engineering education: A big-data bibliometric analysis. *Journal of Engineering Education*, 103(3):486–514, 2014.



TAXONOMÍA DE PALABRAS CLAVE PARA INVESTIGACIÓN SOBRE EDUCACIÓN DE INGENIERÍA

En este anexo se presenta la taxonomía de palabras clave (*keywords*) para la investigación sobre educación de ingeniería (EER, por sus siglas en inglés) creada por Finelli et al. [2015]. Este trabajo ha sido descrito en la Sección 3.5, aquí solo se presenta la taxonomía y la guía de uso.

A.1 TAXONOMÍA DE PALABRAS CLAVE PARA EER

La taxonomía de palabras clave para EER se reproduce en las siguientes figuras (17, 18, 19, 20, 21, 22, 23 y 24) según la última versión consultada en la página web de la iniciativa¹⁰.

¹⁰ URL <http://taxonomy.engin.umich.edu/taxonomy/eer-taxonomy-version-current/> - última consulta 25/07/2016

1. Assessment	2. Design
1.a. Organizational assessment	2.a. Design practice
1.b. Professional licensure	2.a.i. Ideation
1.b.i. Chartered engineer	2.a.ii. Information gathering
1.b.ii. Professional engineer	2.a.iii. Modeling
1. Fundamentals of Engineering exam	1. Physical modeling
1.c. Program evaluation	a. 3D modeling
1.c.i. Accreditation	b. Prototyping
1. ABET	i. Rapid prototyping
2. Criteria	2. Process modeling
1.c.ii. Advisory boards	a. Flowcharting
1.c.iii. Course assessment	2.a.iv. Needs analysis
1.c.iv. External evaluation	2.a.v. Problem definition
1.c.v. Multilevel program assessment	2.a.vi. Product testing
1.d. Student assessment	2.b. Design projects
1.d.i. Assessment tools	2.b.i. Capstone projects [syn: Senior projects, Senior design]
1. Feedback	2.b.ii. Design competitions
a. 360 degree	2.b.iii. Multidisciplinary design
2. Grades	2.c. Design process
a. Automated grading	2.c.i. Human centered design [syn: User centered design]
b. Grading systems	2.c.ii. Product archaeology [syn: Product dissection, Reverse engineering]
c. Inflation	2.c.iii. Product development
3. Concept Inventory	2.d. Design thinking
4. Portfolios	
5. Rubric	
6. Test format [syn: Exam format]	
a. Multiple choice	
b. Open ended tests	
c. Practical examinations [syn: Clinical examinations]	
d. Standardized	
1.d.ii. Knowledge gain	
1.d.iii. Knowledge retention	
1.d.iv. Performance	
1.d.v. Method	
1. Continuous	
2. Diagnostic	
3. Formative	
4. Peer review	
5. Outcomes based assessment	
6. Self assessment	
7. Summative	
1.d.vi. Setting	
1. Individual	
2. Group	
3. Online	
4. Workplace	

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 17: Taxonomía de palabras clave para EER - pág. 1

3. Diversity

- 3.a. Diversity concerns
 - 3.a.i. Bias
 - 3.a.ii. Discrimination
 - 3.a.iii. Inclusivity
 - 3.a.iv. Multiculturalism
 - 3.a.v. Student diversity
 - 3.a.vi. Underrepresentation [syn: Underrepresented students]
 - 3.a.vii. Workplace diversity
- 3.b. Types of diversity
 - 3.b.i. Gender
 - 1. Female [syn: Women, Girls]
 - 2. Male
 - 3. Transgender
 - 3.b.ii. Individual differences
 - 1. Learning styles
 - 2. Personality types
 - 3.b.iii. Nontraditional students
 - 1. Commuter students
 - 2. Part time students
 - 3. Transfer students
 - 4. Veterans
 - 3.b.iv. Race/Ethnicity
 - 3.b.v. Sexual orientation
 - 3.b.vi. Student background
 - 1. First generation
 - 2. International students
 - 3. Socioeconomic status
 - 3.b.vii. Students with disabilities

4. Educational level

- 4.a. Continuing education
- 4.b. Graduate education [syn: Postgraduate]
 - 4.b.i. Graduate
 - 1. Master's students
 - 2. PhD students [syn: Doctoral students]
 - 4.b.ii. Supervision
- 4.c. Higher education [syn: College, University]
- 4.d. P-12 [syn: P12, K-12, K12]
 - 4.d.i. Elementary school [syn: Primary school]
 - 4.d.ii. High school
 - 1. Advanced Placement courses
 - 2. Pre college preparation
 - 4.d.iii. Middle school
 - 4.d.iv. Preschool
 - 4.d.v. Pre-engineering
- 4.e. Postdoctoral studies
- 4.f. Undergraduate
 - 4.f.i. First year [syn: Freshmen, Freshman]
 - 1. First year curriculum
 - 2. First year experience
 - 4.f.ii. Junior
 - 4.f.iii. Senior
 - 4.f.iv. Sophomore

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

5. Educational setting

- 5.a. Engineering curriculum
- 5.b. Engineering fields
 - 5.b.i. Architectural engineering
 - 5.b.ii. Biomedical engineering
 - 5.b.iii. Chemical engineering
 - 5.b.iv. Civil engineering
 - 5.b.v. Computer engineering
 - 5.b.vi. Computer science
 - 5.b.vii. Construction engineering
 - 5.b.viii. Electrical engineering
 - 5.b.ix. Engineering technology
 - 5.b.x. Environmental engineering
 - 5.b.xi. Information technology
 - 1. Green engineering
 - 2. Sustainability
 - 5.b.xii. Industrial engineering
 - 5.b.xiii. Manufacturing
 - 5.b.xiv. Materials science and engineering
 - 5.b.xv. Mechanical engineering
 - 5.b.xvi. Ocean engineering [syn: Marine engineering]
 - 5.b.xvii. Nuclear engineering
- 5.c. Informal learning [syn: Outreach]
- 5.d. Institution type
 - 5.d.i. Baccalaureate institutions
 - 5.d.ii. Community colleges
 - 5.d.iii. Doctoral institutions
 - 5.d.iv. Hispanic serving institutions (HSIs)
 - 5.d.v. Historically black colleges/universities (HBCUs)
 - 5.d.vi. Master's institutions
 - 5.d.vii. Single gender campuses
 - 5.d.viii. Technical colleges
 - 5.d.ix. Tribal colleges
- 5.e. Learning environment
 - 5.e.i. Classroom
 - 5.e.ii. Co-curricular
 - 5.e.iii. Extracurricular
 - 5.e.iv. Honors programs
 - 5.e.v. International programs
 - 5.e.vi. Laboratory
 - 5.e.vii. Learning communities
 - 5.e.viii. Studio
 - 5.e.ix. Undergraduate research

6. Educational technology [syn: E-learning]

- 6.a. Computer-based instruction [syn: Internet-based instruction]
 - 6.a.i. Games
 - 6.a.ii. Educational software
- 6.b. Electronic communication
 - 6.b.i. Blog
 - 6.b.ii. Email
 - 6.b.iii. Groupware
 - 6.b.iv. Instant messaging
 - 6.b.v. Online discussions
 - 1. Web discussions [syn: Chat]
 - 2. Wikis
 - 6.b.vi. Online repositories
 - 6.b.vii. Social media
 - 6.b.viii. Streaming Media
 - 1. Streaming audio [syn: Podcast]
 - 2. Streaming video
- 6.c. Learning technology
 - 6.c.i. Adaptive computer learning
 - 6.c.ii. Learning management systems
 - 6.c.iii. Personal response system [syn: Clicker]
 - 6.c.iv. Simulation
 - 6.c.v. Mobile applications
 - 6.c.vi. Pen and touch devices
 - 6.c.vii. Virtual reality
- 6.d. Learning modality
 - 6.d.i. Blended learning
 - 6.d.ii. Distance learning
 - 1. Asynchronous
 - 2. Massive Open Online Classes (MOOCs)
 - 3. Synchronous
 - 6.d.iii. Remote laboratory [syn: Virtual laboratory]

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 19: Taxonomía de palabras clave para EER - pág. 3

7. Instruction

- 7.a. Conceptual learning [syn: Conceptual change]
 - 7.a.i. Concept inventories
 - 7.a.ii. Concept maps
 - 7.a.iii. Misconceptions
 - 7.a.iv. Preconceptions
 - 7.a.v. Threshold concepts
- 7.b. Faculty [syn: Instructors]
 - 7.b.i. Faculty attitudes
 - 7.b.ii. Faculty development [syn: Educational development]
 - 1. Pedagogical content knowledge
 - 2. Reflective practice
 - 3. Teaching skills
 - 7.b.iii. Instructional role
 - 1. Adjunct
 - 2. Advisor
 - 3. Graduate teaching assistant
 - 4. Instructor
 - 5. Peer teaching assistant
 - 7.b.iv. Teaching philosophies
 - 7.b.v. Team teaching
- 7.c. Institutional change [syn: Institutional transformation, Organizational change]
 - 7.c.i. Evidence-based practice
 - 7.c.ii. Institutional culture
 - 7.c.iii. Instructional change
 - 7.c.iv. Research to practice
 - 1. Adoption
 - 2. Diffusion
 - 3. Dissemination
 - 7.c.v. Theories of change
- 7.d. Instructional design
 - 7.d.i. Alignment
 - 7.d.ii. Bloom's taxonomy
 - 7.d.iii. Course design
 - 7.d.iv. Backwards design
 - 7.d.v. Learning objectives
- 7.e. Instructional methods [syn: Pedagogy]
 - 7.e.i. Active learning
 - 1. Experiential learning
 - 2. Inquiry based learning
 - 3. Peer instruction
 - 4. Challenge based instruction
 - 7.e.ii. Critical pedagogy
 - 7.e.iii. Flipped classroom
 - 7.e.iv. Lecture
 - 7.e.v. Model-eliciting activities
 - 7.e.vi. Mutual learning models
 - 1. Collaborative learning
 - 2. Cooperative learning
 - 3. Team based learning
 - 7.e.vii. Problem based learning
 - 7.e.viii. Project based learning
 - 7.e.ix. Service learning
- 7.f. Teaching evaluations

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 20: Taxonomía de palabras clave para EER - pág. 4

8. Outcomes

-
- 8.a. Communication
 - 8.a.i. Audiences
 - 8.a.ii. Communication skills
 - 1. Nonverbal
 - 2. Verbal
 - a. Listening
 - b. Oral presentations
 - c. Speaking
 - 3. Visual communication
 - a. Engineering graphics
 - b. Illustrations
 - 4. Visualization [syn: Spatial skills]
 - 5. Written communication
 - a. Argumentation
 - b. Reading
 - c. Writing
 - 8.a.iii. Foreign languages
 - 8.a.iv. Technical communication
 - 8.b. Competence
 - 8.c. Computing skills [syn: Computing knowledge]
 - 8.d. Creativity
 - 8.e. Critical thinking
 - 8.f. Engagement
 - 8.g. Engineering standards
 - 8.h. Entrepreneurship
 - 8.i. Ethics
 - 8.i.i. Academic dishonesty [syn: Academic integrity]
 - 1. Plagiarism
 - 8.i.ii. Social justice
 - 8.i.iii. Social responsibility
 - 8.j. Information literacy [syn: Information fluency]
 - 8.k. Innovation
 - 8.l. Intercultural competence [syn: Global]
 - 8.l.i. Cultural schemas
 - 8.m. Leadership
 - 8.n. Lifelong learning
 - 8.o. Problem solving
 - 8.p. Professional skills [syn: Soft skills]
 - 8.q. Scientific literacy
 - 8.r. Student perception
 - 8.s. Student experience
 - 8.t. Teamwork [syn: Team skills]
 - 8.u. Reflection
 - 8.u.i. Critical reflection
 - 8.v. Systems thinking

9. Professional practice

-
- 9.a. Careers
 - 9.a.i. Career choice
 - 9.a.ii. Career paths
 - 9.b. Engineering profession
 - 9.b.i. Employers
 - 9.b.ii. Employment
 - 9.b.iii. Workplace culture
 - 9.c. Engineering management
 - 9.d. Industry involvement
 - 9.d.i. Cooperative education
 - 9.d.ii. Industry sponsorship
 - 9.d.iii. Internships

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

10. Recruitment and retention

- 10.a. Academic support
 - 10.a.i. Supplemental instruction
 - 10.a.ii. Tutoring
- 10.b. Achievement
- 10.c. Advising
 - 10.c.i. Academic advising
 - 10.c.ii. Coaching
 - 10.c.iii. Mentoring
 - 1. Peer mentoring
- 10.d. Preparation
- 10.e. Recruitment
 - 10.e.i. Engineering recruitment
 - 1. Engineering pipeline
 - 10.e.ii. Matriculation
 - 10.e.iii. Enrollment
- 10.f. Retention
 - 10.f.i. Attrition
 - 10.f.ii. Persistence
 - 10.f.iii. Retention rate
 - 10.f.iv. Scholarships
 - 10.f.v. Time to degree
- 10.g. Study behaviors
 - 10.g.i. Study groups
 - 10.g.ii. Time management
- 10.h. Student development
 - 10.h.i. Absenteeism
 - 10.h.ii. Mental health
 - 1. Test anxiety
 - 2. Depression
 - 3. Stress
 - 10.h.iii. Physical health

11. Related fields

- 11.a. Engineering economics
 - 11.a.i. Employability
 - 1. Industry demand
- 11.b. Education policy
 - 11.b.i. Bologna process
 - 11.b.ii. Common core state standards
- 11.c. Mathematics
 - 11.c.i. Calculus
 - 11.c.ii. Complex numbers
 - 11.c.iii. Differential equations
 - 11.c.iv. Engineering mathematics
 - 11.c.v. Graphing
 - 11.c.vi. Linear algebra
 - 11.c.vii. Pre-calculus
 - 11.c.viii. Probability theory
 - 11.c.ix. Statistics
- 11.d. Philosophy of engineering education
- 11.e. Science
 - 11.e.i. Biology
 - 11.e.ii. Chemistry
 - 11.e.iii. Geoscience
 - 11.e.iv. Life science
 - 11.e.v. Physical science
 - 11.e.vi. Physics
 - 11.e.vii. Technology applications
- 11.f. STEM
- 11.g. Technology studies

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 22: Taxonomía de palabras clave para EER - pág. 6

12. Research approaches

- 12.a. Data collection
 - 12.a.i. Analytics
 - 12.a.ii. Focus groups
 - 12.a.iii. Interviews
 - 12.a.iv. Observations
 - 12.a.v. Multi-institution
 - 12.a.vi. Survey
- 12.b. Research ethics
 - 12.b.i. Ethical treatment of subjects
 - 12.b.ii. Professional research ethics
- 12.c. Research evaluation criteria
 - 12.c.i. Credibility
 - 12.c.ii. Dependability
 - 12.c.iii. Generalizability
 - 12.c.iv. Reliability
 - 12.c.v. Transferability
 - 12.c.vi. Trustworthiness
 - 12.c.vii. Validity
- 12.d. Research methods
 - 12.d.i. Design-based research
 - 12.d.ii. Mixed methods research
 - 12.d.iii. Qualitative
 - 1. Case Study
 - 2. Content analysis
 - a. Discourse analysis
 - b. Document analysis
 - 3. Ethnography
 - 4. Grounded theory
 - 5. Phenomenology
 - 6. Phenomenography
 - 12.d.iv. Quantitative
 - 1. Data correlation
 - 2. Descriptive statistics
 - 3. Experimental research
 - 4. Factor analysis
 - 5. Inferential statistics
 - 6. Psychometric analysis
 - 7. Regression
 - 8. Structural equation modeling
 - 12.d.v. Systematic review
 - 1. Meta-analysis

13. Theoretical frameworks

- 13.a. Affective theories
 - 13.a.i. Emotional learning
 - 13.a.ii. Motivation
 - 1. Achievement goal orientation theory [syn: Deep learning, Mastery learning]
 - 2. Attribution theory
 - 3. Behavior theory [syn: Behaviorism]
 - 4. Expectancy Value theory
 - 5. Self-determination theory
 - 13.a.iii. Self efficacy
- 13.b. Cognitive theories
 - 13.b.i. Constructivist
 - 1. Expert-novice
 - 13.b.ii. Knowledge transfer
 - 13.b.iii. Self regulated learning
 - 1. Metacognition
- 13.c. Critical theory
- 13.d. Developmental theory
 - 13.d.i. Adult learning theory
 - 13.d.ii. Agency
 - 13.d.iii. Model of domain learning
 - 13.d.iv. Identity
 - 13.d.v. Perry's model of intellectual development
 - 13.d.vi. Piaget's theory of cognitive development
- 13.e. Social cognitive theories [syn: Social learning theory]
 - 13.e.i. Activity theory
 - 13.e.ii. Cognitive apprenticeship
 - 13.e.iii. Community of practice
 - 13.e.iv. Social cognitive career theory

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 23: Taxonomía de palabras clave para EER - pág. 7

14. Teams [syn: Groups]

- 14.a. Interdisciplinary
- 14.b. Mental models
- 14.c. Multidisciplinary
- 14.d. Self managing work teams
- 14.e. Team dynamics
 - 14.e.i. Nominal group technique
 - 14.e.ii. Team development [syn: Group development]
 - 14.e.iii. Team formation
 - 14.e.iv. Team performance
 - 14.e.v. Team roles
- 14.f. Teamwork training
- 14.g. Transdisciplinary
- 14.h. Virtual teams [syn: Distributed]

This work is licensed under a Creative Commons Attribution-Share Alike 4.0 International License.

Figura 24: Taxonomía de palabras clave para EER - pág. 8

A.2 GUÍA DE USO DE LA TAXONOMÍA

Además de presentar la taxonomía los autores también incluyen una guía para la asignación de palabras clave. A continuación se incluye la explicación de las categorías presentada en la página web de la taxonomía y en la Figura 25 se reproduce la guía incluida en el artículo principal del trabajo de investigación.

A.2.1 *Guidelines for assigning keywords*

Identify one or two keywords from each of the following three categories:

- *Context/ focus / topic. This category involves the context and the background of the paper in terms of the body of knowledge in the field of engineering education. This category also involves the scale of the study (e.g. small group, department, society, multi or single institution), the population (e.g. students in different educational level, faculties/instructors, female/male), field of engineering education (e.g. mathematics, biology, life science, and technology studies), and the stage in education system (e.g. pre-college, graduate, professional development/performances).*
- *Purpose/ target / motivation. This category involves the purpose of the study, the motivation for doing it (e.g. diffusion, adoption), and the outcomes (e.g. communication skills, critical thinking, and scientific literacy).*
- *Research approach. This category involves the underlying theory that guides the work, as well as the research method and tools that are applied for the project (e.g. qualitative/quantitative analysis, interview, survey, focus group).*

Keyword category (select one or two keywords from each of the following three categories)	Related taxonomy branches
Context/focus/topic	1. Assessment
Focus of study	2. Design
Educational level (e.g. pre-college, graduate student, practicing engineer)	3. Diversity
Field of engineering (e.g., physics, civil engineering)	4. Educational level
Population (e.g., undergraduate students, faculty, women)	5. Educational settings
Scale of the study (e.g., individual students, small groups, department, multi- or single-institution)	6. Educational technology
	7. Instruction ^a
	9. Professional practice ^a
	10. Recruitment and retention ^a
	11. Related fields
	13. Theoretical frameworks ^a
	14. Teams
Purpose/target/motivation	7. Instruction ^a
Outcome under study (e.g., communication skills, critical thinking, scientific literacy)	8. Outcomes
Purpose (e.g., diffusion, adoption)	9. Professional practice ^a
	10. Recruitment and retention ^a
	13. Theoretical frameworks ^a
Research approach	12. Research approaches
Underlying theory	13. Theoretical frameworks ^a
Research methods (e.g., qualitative analysis, quantitative analysis)	
Research tools (e.g., interview, survey, focus group)	

^aWords appearing in these taxonomy branches are relevant for multiple keyword categories.

Figura 25: Taxonomía de palabras clave para EER - Guía para selección de palabras clave.

VOCABULARIO CONTROLADO INICIAL SOBRE EDUCACIÓN DE INGENIERÍA DE SOFTWARE

En este anexo se presenta el vocabulario controlado inicial obtenido durante la instancia de trabajo que se describe en el Capítulo 5.

El vocabulario controlado inicial sobre educación de ingeniería de software con sus tres facetas principales se presenta en el Cuadro 12.

Cuadro 12: Vocabulario controlado inicial sobre Educación de ingeniería de software

Term	Synonyms
What to teach	
Communication and collaboration skills	Collaborative and communicative skills, Teamwork skills
Complex systems	
Component-based software engineering	
Modelling techniques	
Object oriented programming design	
Personal Software Process	PSP
Real world problems	
Requirements specification	
Secure Software	
Soft skills	
Software architecture	
Software design	
Software maintenance	
Software metrics	
Software process	
Software process management	
Software project management	
<i>Continúa en la siguiente página</i>	

Cuadro 12 – Continuación de la página anterior

Term	Synonyms
Software quality	
Software reuse	
Software testing	
Technical skills	
Test-driven development	
Unified modelling language	
How to teach	
Teaching approaches and methods	
Academia and industry collaboration	Industry and university partnership
Capstone project	
Case studies	Case study approach
Globally distributed project course	Distributed project group work
Problem based learning	PBL
Real-client projects	Development real-client projects
Simulation of real life	Simulation of real-life environment, Simulation of real-life problems
Software engineering project course	Team project courses
Team projects	Group projects, Team-based projects
Learning and environment materials	
Learning environment	
Software engineering body of knowledge	SWEBOK
Teaching materials	
Teaching software tools	Classroom software tools
Where to teach	
Computer science	Computer science course, Computer science department, Computer science program, Computer science curriculum
Degree in software engineering	Bachelor science in software engineering, BSSE, Degree programme in software engineering, Undergraduate program
Industry training	Training programs in industry

Continúa en la siguiente página

Cuadro 12 – *Continuación de la página anterior*

Term	Synonyms
Information systems	Information systems course, Information systems department
Master degree in software engineering	Graduate curriculum, Graduate
Software engineering course	

CATEGORÍA 'TEACHING APPROACHES AND METHODS' EXPANDIDA

En este anexo se presenta la composición de la categoría '*teaching approaches and methods*' de la faceta '*how to teach*' obtenida durante la instancia de trabajo que se describe en el Capítulo 6.

El detalle de las entradas de la categoría '*teaching approaches and methods*' se muestra en el Cuadro 13.

Cuadro 13: Categoría '*teaching approaches and methods*' de la faceta '*how to teach*' expandida.

Term	Academia and industry collaboration
Definition note	It refers to any effort made together between academia (one or more universities) and industry (one or more companies) for the education of software engineering concepts and practices.
Synonyms	
Facet	How to teach
Related terms	
Related content	

Term	Active learning
Definition note	Active learning is generally defined as any instructional method that engages students in the learning process. In short, active learning requires students to do meaningful learning activities and think about what they are doing.
Synonyms	
Facet	How to teach
Related terms	

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Related content	The definition included here is based in the classic work: Bonwell, C. C., & Eison, J. A. (1991). <i>Active Learning: Creating Excitement in the Classroom</i> . 1991 ASHE-ERIC Higher Education Reports. ERIC Clearinghouse on Higher Education, The George Washington University, One Dupont Circle, Suite 630, Washington, DC 20036-1183.
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Term	Case-based learning
Definition note	Students analyze, draw inferences and make decisions given a case study - a detailed specific situation with a open-ended question or problem (often based on a real life scenario). Case studies can be presented to individuals or groups.
Synonyms	Case studies, CBL, case study approach
Facet	How to teach
Related terms	
Related content	

Term	Clickers
Definition note	A clicker (or a audience response system) is a combination of hardware and software that enables the instructor to pose real-time questions to students during a lecture. The students usually register their responses using handheld clickers, although other input such as laptops may be used. After receivers transmit the responses to the instructor's workstation, the software compiles and displays the results.
Synonyms	Audience response system, personal response system, classroom response system, student response system
Facet	How to teach
Related terms	
Related content	In one of the most cited paper Caldwell includes a comprehensive overview of the technique and best-practice tips: Caldwell, J.E. (2007) <i>Clickers in the Large Classroom: Current Research and Best-Practice Tips</i> , CBE-Life Sciences Education, vol. 6, n.º 1, pp. 9-20.

Term	Concept mapping
------	-----------------

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Definition note	Concepts maps can be defined as a knowledge representation language. In short, the students create or use graphic structures that arrange key ideas or concepts in a hierarchical set of nodes with lines or arrows that indicate linkages and relationships between them.
Synonyms	
Facet	How to teach
Related terms	
Related content	Concept mapping were developed by Joseph Novak. In this two often cited works he presents the basis of concept mapping and Vee diagrams: Novak, J. D. & Gowin, D. B. (1984) <i>Learning How to Learn</i> , Cambridge University Press; and Novak, J. D. (1990) <i>Concept maps and Vee diagrams: two metacognitive tools to facilitate meaningful learning</i> , <i>Instructional Science</i> , Volume 19, Number 1, Page 29.

Term	Cooperative learning
Definition note	Refers to any of a variety of teaching methods in which students work in small groups to help one another learn academic content. Most experts agree that cooperative learning has several components that distinguish it from other small group learning methods. These components may include: positive interdependence (a positive correlation between the gains of individuals and the gains of teams, individual accountability (although learning activities rely on cooperative efforts, individuals are ultimately responsible of their own learning), group processing (group members discuss their progress towards the achievement of their goals and the maintenance of effective working relations), face-to-face interaction (the size of the groups must be small), social and cooperative skills (that must be taught and motivated by the instructor) and appropriate grouping (some authors recommend heterogeneous teams, reflecting varied learning abilities, ethnic and linguistic diversity and a mixture of the sexes).
Synonyms	
Facet	How to teach
Related terms	

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Related content	Three important and cited books about this topic are: Slavin, R. E. (1995) <i>Cooperative Learning: Theory, Research, and Practice</i> , Allyn and Bacon.; Kagan, S. (1994) <i>Cooperative learning</i> , San Clemente, CA: Kagan.; and Johnson, D. W., Johnson, R. T. & Holubec, E. J. (1993) <i>Circles of learning: cooperation in the classroom</i> , 4th ed. Edina, Minn: Interaction Book Co.
-----------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Term	Distance Learning
Definition note	Any educational or learning process or system in which the teacher and instructor are separated geographically or in time from his or her students; or in which students are separated from other students or educational resources. Contemporary distance learning is affected through the implementation of computer and electronics technology to connect teacher and student in either real or delayed time or on an as-needed basis.
Synonyms	
Facet	How to teach
Related terms	
Related content	The selected description is the broader found, taken from: Illyefalvi-Vitez, Z. and Gordon, P. (2004) <i>Distance learning - how to use this new didactic method in education of electronics engineering?</i> , in Electronic Components and Technology Conference. Proceedings. 54th, 2004, vol. 2, pp. 1725-1730 Vol.2.

Term	E-learning
Definition note	E-learning can be defined as instruction delivered electronically via the internet, intranet, or multimedia platforms such as cd-rom or dvd. E-learning is used to describe a wide set of applications and processes, such as web-based learning, virtual classrooms, and digital collaboration.
Synonyms	
Facet	How to teach
Related terms	

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Related content	The definition included here is based on two works: Smart, K. L., & Cappel, J. J. (2006). <i>Students' perceptions of online learning: A comparative study</i> Journal of Information Technology Education, 5 , 201-202.; and Kaplan-Leiserson, E. <i>E-Learning Glossary</i> . (www.learningcircuits.org/glossary.html)
-----------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Term	Game-based learning
Definition note	This technique deals with games that have defined learning outcomes. A game can be defined as an activity that is voluntary and enjoyable, separate from the real world, uncertain, unproductive in that the activity does not produce any goods of external value, and governed by rules.
Synonyms	Games
Facet	How to teach
Related terms	Simulation and Games
Related content	The definition of a game is taken from: Caillois, R. (1961). <i>Man, play, and games</i> . New York: Free Press.

Term	Globally distributed project course
Definition note	It refers to some initiatives within Project-Based Learning. In these cases, students face projects distributed between two or more universities which work collaboratively on this initiative.
Synonyms	
Facet	How to teach
Related terms	Project-Based Learning
Related content	

Term	Interactive lecture demonstrations
Definition note	Students are asked to predict individually the outcome of a classroom demonstration. Later the students interact in small groups, discussing their predictions and explaining their reasoning. Finally, the demonstration is performed and the students discuss and reflect on the results.
Synonyms	

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Facet	How to teach
Related terms	The approach was developed by Sokoloff and Thornton and can be found at: Sokoloff, D. R. & Thornton, R. K. <i>Using interactive lecture demonstrations to create an active learning environment</i> (1997) <i>The Physics Teacher</i> , vol. 35, pp. 340-347; and Sokoloff, D. R. & Thornton, R. K. (2006) <i>Interactive Lecture Demonstrations, Active Learning in Introductory Physics</i> , Wiley.
Related content	

Term	Just-in-time teaching
Definition note	This approach is based on the feedback loop between the students and the instructor. The instructor use the internet to post course materials and warm up assignments before class, and the students use materials to prepare for each class. The instructor uses students responses to enhance the classroom component.
Synonyms	Just-in-time, JiTT
Facet	How to teach
Related terms	
Related content	JiTT was first presented in the book: Novak, G. M., Gavrin, A., Christian, W., & Patterson, E. (1999). <i>Just-in-Time Teaching: Blending Active Learning with Web Technology</i> (1st ed.). Upper Saddle River, NJ, USA: Prentice Hall PTR.

Term	Lectures
Definition note	Lectures are (more or less) continuous expositions of a theme by a speaker to a largely passive recipient audience.
Synonyms	
Facet	How to teach
Related terms	
Related content	The definition included here is based on the classic work: Bligh, D.A. (1972). <i>What's the use of lectures?</i> Harmondsworth, England: Penguin. A new version of this book is available as: Bligh, D. A. (2000) <i>What's the use of Lectures</i> London/San Francisco.

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Term	One-minute papers
Definition note	The instructor asks the students (often in the last minutes of class) to write a quick response to one or more questions regarding the content of the class (typically a lecture). Questions might include: what is the most important thing you learned today? what is the muddiest point still remaining at the conclusion of today's class? After collecting the responses, the instructor reads the answers and ideally responds to them in the next class, or privately on an individual basis.
Synonyms	
Facet	How to teach
Related terms	
Related content	More information of this technique can be found in two often cited works: Cross, K. P. & Angelo, T. A. (1988) <i>Classroom assessment techniques: a handbook for faculty</i> National Center for Research to Improve Postsecondary Teaching and Learning; and Chizmar, J. F., & Ostrosky, A. L. (1998). <i>The One-Minute Paper: Some Empirical Findings</i> . <i>The Journal of Economic Education</i> , 29(1), 3-10.

Term	Problem-based learning
Definition note	Students work in groups to analyze and solve a problematic situation, usually a realistic scenario without a single correct answer, under the supervision of a tutor.
Synonyms	PBL
Facet	How to teach
Related terms	
Related content	Two relevant and cited articles about this topic are: Barrows, H. S. (1986). <i>A taxonomy of problem-based learning methods</i> . <i>Medical education</i> , 20(6), 481-486; and Schmidt, H. G. (1993). <i>Foundations of problem-based learning: some explanatory notes</i> . <i>Medical Education</i> , 27(5), 422-432.

Term	Project-based learning
------	------------------------

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Definition note	Students, typically organized in groups, face open multidisciplinary projects with the instructor playing the role of facilitator or coach. The projects engage students in authentic real-world problems and usually leads to the production of a final product (a design, a mode, a software product, etc).
Synonyms	PjBL
Facet	How to teach
Related terms	
Related content	

Term	Real-client projects
Definition note	It refers to instances of Project-Based Learning with real clients (this means the clients are no teachers or other students, and usually are industry members).
Synonyms	
Facet	How to teach
Related terms	Project-based learning
Related content	

Term	Research-based learning
Definition note	It refers to initiatives that connect teaching with research, which allow partial or full inclusion of students in an investigation based on scientific methods, under the supervision of instructor.
Synonyms	
Facet	How to teach
Related terms	
Related content	

Term	Service learning
Definition note	A form of experiential learning in which students engage in activities that address human and community needs while allowing students to reflect on their service to gain further understanding of course concepts.

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Synonyms	
Facet	How to teach
Related terms	
Related content	

Term	Simulation-based learning
Definition note	This technique deals with simulations that have defined learning outcomes. Simulation is a technique to replace or amplify real experiences with guided experiences, often immersive in nature, that evoke or replicate substantial aspects of the real world in a fully interactive fashion.
Synonyms	
Facet	How to teach
Related terms	Simulation and Games
Related content	The selected description is the broader found, taken from: Gaba, D. M. (2004). <i>The future vision of simulation in health care</i> . <i>Quality and Safety in Health Care</i> , 13(suppl 1), i2-i10.

Term	Simulation and Games
Definition note	Both techniques seeks the instruction by guided experiences in ruled environments (usually immersive). In other words, both have some underlying model, allowable actions that the learner can take, and constraints under which these actions should occur. Additionally, learners observe their actions' consequences. The key distinction is that simulations propose to represent reality and games do not.
Synonyms	
Facet	How to teach
Related terms	
Related content	The definition included here is based on: Garris, R., Ahlers, R., & Driskell, J. E. (2002). <i>Games, Motivation, and Learning: A Research and Practice Model</i> . <i>Simulation & Gaming</i> , 33(4), 441-467.

Term	Software engineering project course
------	-------------------------------------

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Definition note	It refers to initiatives within Project-Based Learning related to specific courses in which students work on a software engineering project (generally involving software development).
Synonyms	
Facet	How to teach
Related terms	Project-Based Learning
Related content	

Term	Student-centered learning
Definition note	Ways of thinking about teaching and learning that emphasise student responsibility and activity in learning rather than what teachers are doing. The students exert a degree of influence over both the content of the course and the learning methods.
Synonyms	
Facet	How to teach
Related terms	
Related content	The first part of the definition included here is taken from: Cannon, R., & Newble, D. (2000). <i>A Handbook for Teachers in Universities and Colleges: A Guide to Improving Teaching Methods</i> . Routledge. Two relevant and cited articles about this topic are: Lea, S., Stephenson, D., & Troy, J. (2003). <i>Higher education students' attitudes to student-centred learning: Beyond 'educational boulimia'?</i> <i>Studies in Higher Education</i> , 28(3), 321-334; and Hannafin, M., Hill, J., & Land, S. (1997). <i>Student-centered learning and interactive multimedia: Status, issues, and implications</i> . <i>Contemporary Education</i> , 68(2), 94-99.

Term	Technology enhanced learning
Definition note	All approaches in which technology is used to support the learning or teaching process.
Synonyms	
Facet	How to teach
Related terms	

Continúa en la siguiente página

Cuadro 13 – Continuación de la página anterior

Related content	The definition included here is taken from: Schweighofer, P., & Ebner, M. (2015). <i>Aspects to Be Considered when Implementing Technology-Enhanced Learning Approaches: A Literature Review</i> Future Internet, 7(1), 26.
-----------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Term	Think-pair-share
Definition note	Students are thought to listen a question, think about the question, to discuss the question in pairs, and finally to share with the total group.
Synonyms	
Facet	How to teach
Related terms	
Related content	This technique was first proposed by Frank Lyman in: Lyman, F. T. (1981). <i>The responsive classroom discussion: The inclusion of all students</i> . In A. S. Anderson (Ed.), <i>Mainstreaming Digest</i> (pp. 109-113). College Park, MD: University of Maryland Press.

Term	Tutorials
Definition note	It is an activity in which the instructor works with one or a small groups of students and that is characterized as a space for discussion. Usually it served to complement other teaching techniques (e.g. lectures) and can be enhanced if the students have done some relevant prior work. It is considered a technique within the student-centered learning approach.
Synonyms	
Facet	How to teach
Related terms	
Related content	Tutorial teaching is part of the learning system at the University of Oxford and involves some particular features. Palfreyman provide a good review of this technique: Palfreyman, D. (Ed.) (2001) <i>The Oxford tutorial: 'thanks, you taught me how to think'</i> (Oxford, Oxford Centre for Higher Education Policy Studies).



INSECTES : 1. Platypria echidna (grossi). — 2. Sagra Ruqueti. — 3. Desmonota variolosa. — 4. Pseclioeris Harwicki. — 5. Catacanthus anehorago. — 6. Callidea signata. — 7. Eulyes melanoptera. — 8. Holco-
 meria spinosa. — 9. Antioecia bilineata. — 10. Cheirochela Fean. — 11. Hucchys incornata. — 12. Gosna frativa. — 13. Aphan novemmaculata. — 14. Hotinus Delesserti. — 15. Phromia rubra. —
 16. Tettigonia pulchella (grossi). — 17. Tettigonia colorata (grossi). — 18. Tettigonia fasciata (grossi). — 19. Phyllomorpha alicia. — 20. Ptilocerus fuscus (grossi). — 21. Carcinocoris Castetii (grossi).
 — 22. Philonthochila (grossi). — 23. Hemiptycha (grossi). — 24. Sphongophorus ballista (grossi). — 25. Sphongophorus Benneti (grossi). — 26. Boecidium globularis (grossi). — 27. Cyphonia furcata (grossi).
 — 28. Cyphonia tritida (grossi). — 29. Heteronotus vulnerans femelle (grossi). — 30. Heteronotus vulnerans male (grossi). — 31. (Ela indata (grossi). — 32. Membracis C. album (grossi). — 33. Combophora
 Beaki (grossi). — 34. Acanthopus splendidus. — 35. Xylocopa caerulea. — 36. Scolia prouer. — 37. Sphex Latreillei. — 38. Synagris cornuta. — 39. Eumenis arcuata. — 40. Pepsis aureipennis. — 41. Mutilla
 cephalotes. — 42. Chlorion lobatum. — 43. Pelecinus polycerator. — 44. Megiscus ruficeps. — 45. Chrysis semicincta (grossi). — 46. Cleptes orientalis (grossi). — 47. Eucherus purpuratus (grossi). —
 48. Stilpnus salina (grossi). — 49. Anthus (grossi). — 50. Anthus (grossi).