

Introducción a la Ingeniería del Conocimiento

Francisco J. Martín Mateos

Dpto. Ciencias de la Computación e Inteligencia Artificial
Universidad de Sevilla

¿Qué es la Ingeniería del Conocimiento (KE)?

- La Ingeniería del Conocimiento (KE) es una disciplina cuyo objetivo es integrar conocimiento en sistemas informáticos para resolver problemas complejos que normalmente requieren un alto nivel de conocimiento experto humano.
- Concretamente, se trata de aplicar metodologías para la construcción, mantenimiento y desarrollo de sistemas basados en conocimiento.

- Identificación de problemas.
- Adquisición del conocimiento.
- Representación del conocimiento.
- Desarrollo de un entorno de trabajo.
- Implementación de las tareas de conocimiento (inferencia, análisis, diagnóstico, explicación, entrenamiento, ...) en un sistema basado en conocimiento (KBS).
- Prueba y validación del KBS.
- Integración y mantenimiento del KBS.
- Revisión y evaluación del KBS.

¿Qué es un Sistema Basado en Conocimiento (KBS)?

- Los sistemas basados en conocimiento (KBS) son herramientas de inteligencia artificial utilizadas en un dominio específico para proporcionar decisiones inteligentes con justificación.
- Los KBS son sistemas basados en métodos y técnicas de Inteligencia Artificial. Sus componentes clave son:
 - Base de conocimiento.
 - Mecanismos de adquisición.
 - Mecanismos de inferencia.

¿Qué es la Ingeniería del Conocimiento (KE)?

- Es una disciplina basada en tres pilares:
 - Herramientas de Inteligencia Artificial: sistemas de producción, redes bayesianas, razonamiento basado en casos, redes neuronales, algoritmos genéticos y evolutivos, algoritmos de aprendizaje, ...
 - Aplicaciones de Conocimiento Experto: sistemas expertos, sistemas de diagnóstico, sistemas de análisis de requisitos, sistemas de predicción, ...
 - Metodologías de Ingeniería del Software: análisis de aplicabilidad, análisis de requisitos, metodologías de desarrollo, implantación, mantenimiento, ...

- Un sistema de gestión de reglas de negocio (BRMS) es una aplicación utilizada para definir, implementar, ejecutar, monitorizar y mantener la variedad y complejidad de las reglas de negocio de una organización o empresa.
- Un BRMS consiste en un conjunto de reglas que describen la normativa funcional de la organización o empresa, un conjunto de herramientas para analizar, visualizar, editar y validar las reglas y un entorno de ejecución en el que se aplican las reglas mediante un motor de inferencia (Business Rules Engine BRE).

- El conocimiento es la asociación funcional explícita entre elementos de información y/o datos.
- Datos, información y conocimiento representan tres niveles distintos y progresivos:
 - Dato: Una unidad elemental de comunicación.
 - Información: Semántica asociada a los datos.
 - Conocimiento: Capacidad de utilizar provechosamente la información.
- Ejemplo: Resolución de un Sudoku.
- Ejercicio: Identificar los tres niveles en otros contextos.

- Redes neuronales: Son herramientas software diseñadas para estimar relaciones entre datos. Las relaciones estimadas son funciones que dependen de una representación numérica de los datos y sus características.
 - Se utilizan en reconocimiento de patrones, filtrado, compresión de datos, control adaptativo, modelado de funciones complejas, ...
 - Es una herramienta de aprendizaje supervisado que no utiliza el conocimiento ni la información asociada a los datos, sólo detectan relaciones funcionales numéricas entre los datos.

- Algoritmos genéticos: Es un modelo de aprendizaje automático en el que se buscan soluciones a problemas complejos mediante un proceso “evolutivo” que consiste en realizar mejoras sucesivas sobre un conjunto de candidatos.
 - Se utilizan en problemas de planificación, logística, diseño de circuitos, ...
 - Los candidatos son codificaciones de los datos y la información asociada a los datos se utiliza para discriminar dichos candidatos. No hacen uso explícito del conocimiento funcional asociado a los datos.

- Minería de datos: Es el proceso de extracción automática de información predictiva oculta en colecciones de datos.
 - Se utilizan para caracterizar datos, discriminar datos, analizar asociaciones y clasificaciones, realizar predicciones, ...
 - Es una herramienta de extracción de información a partir de los datos. El análisis de esta información proporciona nuevas fuentes de conocimiento.

- Redes bayesianas: Es un modelo estadístico que establece relaciones entre datos. Permite analizar el grado de dependencia entre datos procedentes de diferentes fuentes.
 - Se utilizan para categorizar soluciones, procesos de clasificación y diagnóstico, ...
 - Permiten organizar datos, haciendo explícitas las relaciones entre ellos y los grados de dependencia, de esta forma capturan la información relevante asociada a los datos.
 - Formalizan el conocimiento asociado a la relaciones causa/efecto entre los datos.

- Sistemas de producción: Es un modelo computacional basado en la aplicación sucesiva de reglas de producción que hacen explícitas las relaciones entre conjuntos de datos.
 - Se utilizan para la implementación de sistemas expertos en general.
 - Permiten formalizar los datos y la información asociada en forma de hechos; y el conocimiento sobre como utilizar dicha información mediante reglas de producción.

- Kendal, S. y Creen, M.
“An Introduction to Knowledge Engineering”, Springer-Verlag,
2007.