

Tema-2-Adquisicion-del-Conocimie...



ParmigianoReg



Ingenieria del Conocimiento



3º Grado en Ingeniería Informática



Escuela Técnica Superior de Ingenierías Informática y de Telecomunicación
Universidad de Granada

NEW

WUOLAH Print

Lo que faltaba en Wuolah



Imprimir



Tema 2 - Adquisición del Conocimiento

Profesor Juan Luis Castro Peña, Curso 20-21

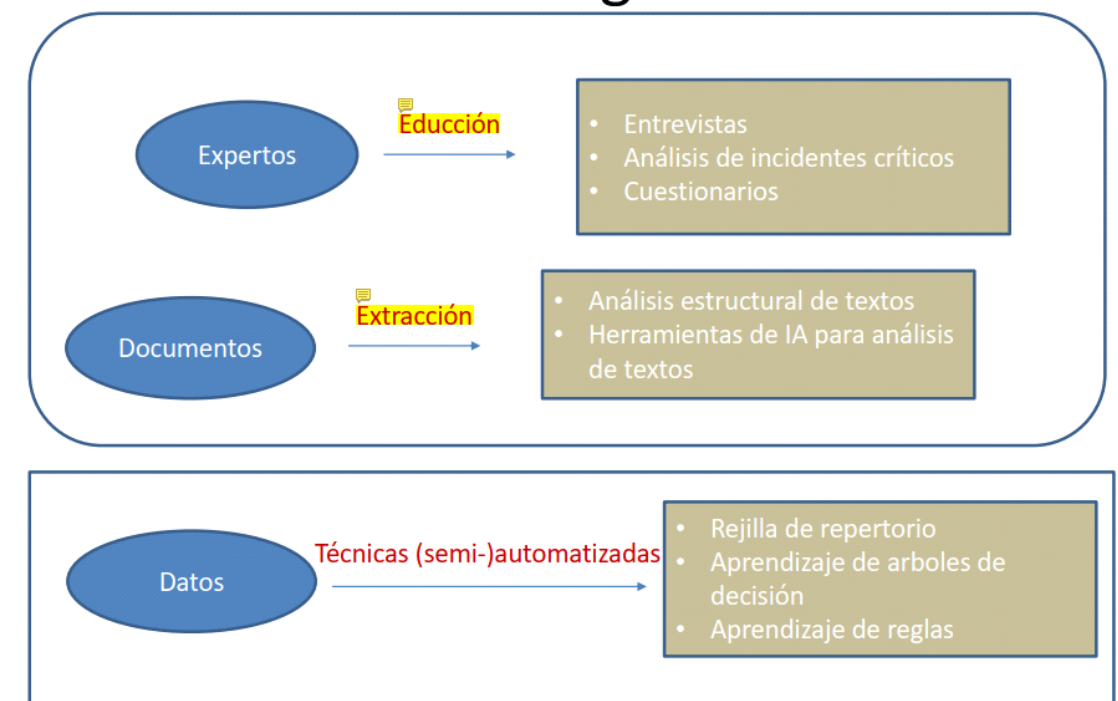
Introducción:

- Adquisición del conocimiento: Es el proceso de recolección de información de cualquier fuente existente (Expertos, libros, revistas, informes, etc.) necesaria para construir un Sistema Basado en Conocimiento. Esto se realiza por lo general en el desarrollo del problema.
 - **Objetivo:** Proporcionar a cada etapa la información que se requiere en cada momento del desarrollo.
 - **Cuando:** No es un paso en concreto en la metodología del desarrollo de un SBC, si no, más bien una tarea que se produce en paralelo a todas las etapas de construcción de estos sistemas (Identificación, Conceptualización, Formalización, Validación y Mantenimiento)
 - Este proceso forma parte de cada fase del desarrollo de un Sistema Basado en Conocimiento.

Situación actual

- La adquisición de conocimientos es la tarea más importante del desarrollo de un SBC.
- Se pueden aplicar algunas técnicas propias de IA pero como herramientas.
- No existe un método automatizado de Adquisición del Conocimiento.
- Existen unas metodologías utilizadas para adquirir correctamente el conocimiento.

En general...



Fuentes de Conocimiento: Personas

- Pueden ser los Expertos, Directivos y Usuarios Finales
 - **Expertos:** Poseen la mayor parte de los conocimientos.
 - **Directivos:** Poseen los objetivos del proyecto, el alcance del sistema, el contexto donde será instalado.
 - **Usuarios:** Comprender el tipo de persona que interactuará con el sistema, sus necesidades y requisitos.

Fuentes de Conocimiento: Documentos

- **Resultados de estudios e investigaciones:** datos, estudios, informes, resultados estadísticos, etc.
- **Publicaciones especializadas:** Contienen la información más al día.
- **Material utilizado para la formación:** Suele contener conocimiento expuestos de manera especialmente clara.
- **Casos:** Algunas empresas suelen registrar los casos que se les presentan en forma de órdenes de reparación, fichas de clientes o pacientes, estudios o almacenamiento de casos.

Fuentes de Conocimiento: Experiencia

- Observación (Visitas): útil para clarificar ideas y entender el esquema general del proceso. Es mejor que explicitar el conocimiento.
- Documentos informales:
 - Notas manuscritas, ayudas de trabajo que circulan en organizaciones.
 - Proporciona a menudo conocimiento heurístico para resolver problemas.
 - Suelen reflejar la experiencia de profesionales de la organización a la hora de enfrentarse a ciertos problemas.
 - Proporciona conocimiento semipúblico.

Extracción y Educción:

- **Extracción de conocimiento:** cuando la fuente de conocimiento se presenta de forma escrita o registrada.
- **Educción de conocimiento:** cuando los conocimientos se obtienen de humanos.

Proceso de adquisición del conocimiento

- **Estudio del problema:** Primeras reuniones y evaluación de viabilidad (De expertos, directivos y usuarios)
- **Extracción de conocimientos** de la documentación
- **Educción de conocimientos** del experto
 - Interrogatorio inicial
 - Investigación profunda

Estudio del Problema

- Consiste en determinar los requisitos funcionales del sistema, o en su caso, las necesidades de los usuarios del futuro sistema o lo que los usuarios esperan del mismo.
- Consiste en definir el problema formalmente, se define como obtener unos valores de salida para unos valores de entrada por medio de una función.
- Se debe introducir al Ingeniero de Conocimiento en el dominio a un nivel tal que sea capaz de desarrollar un estudio de viabilidad del Sistema Experto donde se determine si la tarea del experto es tratable o no, mediante la Ingeniería del Conocimiento.
 - Se genera el esquema general del proceso.

Estudio del Problema: Primeras reuniones

- Se buscan conocimientos generales, no de detalle, así como familiarizarse con la terminología del dominio. El ingeniero debe centrarse en el entorno de la tarea y los usuarios.
- Hay que dirigir la entrevista a definir el problema: entradas, salidas e interacción del sistema.
- Posteriormente hay que conseguir el esquema general del proceso de razonamiento en forma de tareas, subtareas y las relaciones y dependencias entre ellas.
- Hay que analizar la viabilidad de obtener los hechos en que se basa el razonamiento.
- Hay que evitar a toda costa que nuestras opiniones e ideas interfieran en los resultados (Error muy común que ralentiza esta fase)
- Una o dos reuniones ha de ser suficiente. Es un fallo muy típico hacer más.

Estudio del Problema: Análisis de Viabilidad

- Una vez realizadas las reuniones, se deben poder responder las siguientes preguntas.
- ¿Es factible obtener los datos que se utilizan de entrada? ¿Cuánto coste tendría?
 - Pueden también ser introducidos por el usuario si calcularlos es demasiado complejo o costoso.

NEW

WUOLAH Print

Lo que faltaba en Wuolah



Imprimir



- ☐ Todos los apuntes que necesitas están aquí
- ☐ Al mejor precio del mercado, desde **2 cent.**
- ☐ Recoge los apuntes en tu copistería más cercana o recíbelos en tu casa
- ☒ Todas las anteriores son correctas



- ¿El esquema general del proceso de razonamiento se puede representar?
 - La solución más común es un sistema modular de tareas.
- ¿Cada tarea podría representarse como un SBC simple mediante un modelo de representación del conocimiento básico?
 - Habría que reiterar el proceso para cada tarea que no resulte básica.

Extracción del Conocimiento

- Aprender sobre el dominio tanto como sea posible antes de comenzar las sesiones con el experto.
- Se pretende evitar, o al menos, reducir el tiempo que en otro modo el experto utilizaría para poner al tanto al Ingeniero de Conocimiento en el tema.

Extracción del Conocimiento: Consulta con el Experto

- Se le pide al experto que indique los documentos más importantes a consultar, el material más relevante de la colección de manuales existentes y que explique la terminología utilizada y conceptos discutidos en la documentación.
- El experto debe proporcionar detalles que estén omitidos en la documentación pero que son cruciales para el entendimiento global de la tarea.
- Señalar los puntos donde la práctica real difiere de los procedimientos documentados.
- Explicar las anotaciones a mano hechas sobre los documentos de trabajo.

Extracción del Conocimiento: Análisis estructural de textos

- Ante un determinado texto, se establece una serie de términos que deben ser buscados en el texto.
- Los términos a buscar están preestablecidos por la técnica y son dependientes del dominio.
 - Por ejemplo Enfermedad, Medicina, Terapia, Tratamiento...
- Los términos a buscar están preestablecidos pero son independientes del dominio.
 - Por ejemplo X se define como Y; X está relacionado con Y.

Análisis Estructural de Textos: Términos independientes del dominio

- Los términos que son independientes del dominio son conceptos y relaciones.
- Se deben buscar estructuras tipo "definición" para los conceptos y estructuras tipo "afirmación relacional" para relaciones.
- Permite extraer conocimiento sin entender por completo el tema.
- Se necesita:
 - Las estructuras a buscar y el tipo de conocimiento que se obtendrá de cada una, sea un concepto, relación, característica o valor.
 - Un modo de detectar esas estructuras.

Tipos de Estructuras a Buscar

- **Definiciones:** Introducción de un concepto nuevo en el texto.
- **Afirmaciones:** Una frase que establece una verdad, se buscan aquellas afirmaciones que expresan relaciones entre conceptos
- **Leyes:** Las leyes de un dominio establecen sus principios básicos, así como las reglas que fijan el funcionamiento de objetos en el dominio.
- **Procedimientos:** Los procedimientos de un dominio establecen los pasos para la resolución de problemas en el dominio. Al igual que el caso anterior, los conocimientos proporcionados por esta estructura están más allá del objetivo de la extracción de conocimientos a partir de la documentación.

Modo de Detectar Estructuras

- Búsqueda de patrones
 - Definiciones: A se usa para B, A es un B que C, A está compuesto de B y C, etc.
 - Afirmaciones: A es causa de B, A se relaciona con B, A es finalidad de B, etc.
- Buscar títulos de secciones, subsecciones y textos en **negrita** o *cursiva*.
- El trabajo es simple y tedioso
 - Hojear el texto deteniéndose solo en las definiciones y afirmaciones de relación, estas frases se señalan.
 - Analizar las frases señaladas para extraer los conocimientos buscados: conceptos, relaciones y definiciones de conceptos.

- Existen herramientas para esta tarea, como **Word2Vec**. Se codifican en vectores las palabras de un texto y las palabras cercanas tienen significados similares.

Educción del Conocimiento

- El Ingeniero de Conocimiento deberá marcar la perspectiva y profundidad deseada en cada sesión.
- Dada la complejidad intrínseca del proceso, cada sesión debe ser controlada en muchos otros aspectos.

Ciclo de Educación

- Preparación de la sesión
 - Información a tratar, amplitud o profundidad de cada una, técnica adecuada y preparación de preguntas.
- Sesión
 - Repaso del análisis de la última sesión, explicación al experto de los objetivos de la nueva sesión, educación y comentarios del experto.
- Transcripción: Se debe escribir la entrevista mas que grabarla.
- Análisis de la sesión
 - Lectura para obtener una visión general.
 - Extraer conocimientos concretos: Subdominio, tareas, problemas, conceptos y relaciones.
 - Lectura para recuperar detalles olvidados.
 - Crítica para mejorar por parte del Ing. de Conocimiento: problemas que ha habido, información faltante, preguntas que se debieron haber realizado.
- Evaluación
 - Las preguntas a responder en este momento son:
 - ¿Se han conseguido los objetivos?
 - ¿Es necesario volver sobre el mismo objetivo?
 - Número y tipo de sesiones necesarias para cubrir el área.
 - Esta fase del ciclo de educación se funde con la primera fase del siguiente ciclo, evaluando si se han conseguido los objetivos, si se necesitan más sesiones responden a las preguntas que se formulan en la primera fase del ciclo.

Técnicas de Educación

- **Métodos Directos:** Se le pregunta directamente al experto que es lo que sabe, el experto es la única fuente de información y el IC confía plenamente en lo que dice el experto.
 - "Dime lo que sabes"
 - Aquí entran entrevistas y cuestionarios.
- **Métodos Indirectos:** Se utilizan porque no siempre los expertos pueden acceder a los detalles de sus conocimientos o procesos mentales, y para confirmar lo adquirido con técnicas directas.
 - "Déjame comprobar lo que sabes"
- **Otras técnicas**
 - Entrevistas (Abiertas o Estructuradas)
 - Cuestionarios
 - Observación de tareas habituales
 - Incidentes críticos
 - Clasificación de conceptos
 - Análisis de protocolos
 - Emparrillado
 - Inducción

Entrevista Abierta

- El IC plantea de manera espontánea las preguntas al experto. Esto no quiere decir que esta técnica no necesite de planificación y control.
- Como cualquier sesión de educación, el IC habrá fijado el tema o perspectiva a tratar con el experto así como la profundidad de los conocimientos a educir.
- La falta de conocimientos sobre la perspectiva fijada y el requisito de un grano grueso en el tema lleva al IC a utilizar este tipo de entrevista como la técnica más adecuada en una determinada sesión de educación.

Preguntas típicas de una Entrevista Abierta

Iniciales

- ¿Cómo resuelves este problema?
- ¿Cuáles son los elementos que influyen cuando resuelves el problema?
- ¿Cuáles son las informaciones que necesitas antes de empezar el tratamiento del problema?

Alentar

- ¿Qué haces a continuación?
- ¿Puedes describir lo que quieres decir con eso?
- ¿Por qué haces eso?

Alerta

- ¿En qué se parece y diferencia este problema con los típicos del dominio?
- ¿Qué tipos de datos necesita el problema?
- ¿Qué clase de soluciones son adecuadas para el problema?
- ¿Qué constituye una explicación o justificación adecuada de la solución del problema?

Entrevista Estructurada

- El IC planifica de antemano las preguntas que debe realizar al experto durante la sesión.
- Debe formular y agrupar las cuestiones lógicamente. Por lo general los grupos conciernen a acciones o procesos que se han identificado en sesiones previas.
- El IC plantea un grupo de cuestiones sobre determinado objeto o atributo y una vez resueltas pasa al grupo siguiente.
- Las preguntas a plantear deben centrarse en los conocimientos del experto sobre los conceptos, relaciones e inferencias del experto.

Preguntas típicas en una Entrevista Estructurada

- ¿Qué tipo de cosas te gustaría saber acerca del problema cuando empiezas a sopesarlo?
- ¿Qué hechos o hipótesis intentas establecer cuando piensas sobre el problema?
- ¿Cuáles son los factores que influyen en la forma en que razonas en el problema?
- ¿Qué tipo de valores puede tener ese objeto?, ¿qué rango de valores está permitido?
- ¿Este valor depende de otros factores? En caso afirmativo, ¿cuáles son?
- ¿Es este factor necesario para resolver todos los problemas en el dominio, o sólo para algunos? En este caso, ¿cuáles?

- ¿Podrías explicarme este concepto en mayor detalle?
- ¿Qué ocurre en este punto?
- ¿Por qué te planteaste este problema?
- ¿Es correcta esta secuencia?
- ¿Es esto lo que haces en esta situación?
- ¿Te parece que este diagrama muestra correctamente el orden de tus decisiones?
- ¿Están incluidos aquí todos los conceptos relacionados con tal tema?

Limitaciones de las entrevistas

- Puntos fuertes
 - Sirven para compilar los conceptos básicos de la tarea.
 - Sirven para obtener información conceptual implicada en el problema.
 - Sirven para extraer conocimientos de relaciones, valores y acciones.
- Puntos débiles

- Las entrevistas consumen mucho tiempo, sobre todo para el Ingeniero de Conocimiento.
- Confían en la memoria del experto.
- Existen problemas con el lenguaje.
 - Eliminación de componentes clave en un proceso de razonamiento.
 - Palabras con referencias no especificadas.
 - Referencias hechas con comparativas.
 - Condensación, con palabras, de procesos complejos.
 - Implicación de una conexión causal entre eventos.

Observación de tareas habituales

- Con frecuencia, la mejor forma de descubrir cómo hace un juicio un experto es observar al experto trabajar en una problemática real habitual.
- Una primera decisión a tomar es como registrar las acciones del experto. Se puede sencillamente observar, tomar notas e intentar seguir al mismo tiempo el proceso de pensamiento del experto. Otra alternativa es grabarlo y posteriormente revisarlo con el experto.
- El IC no interfiere en la actuación del experto en la solución de sus tareas reales cotidianas.
- Puede ser también el entrenamiento de un novato.
- Ventajas
 - Proporciona al IC una primera idea de los tipos de conocimientos y habilidades implicados en el dominio
 - Proporciona conocimientos básicos del dominio y ayuda a que el IC comprenda la tarea del experto.
 - Es útil para captar conocimientos procedimentales o para entender características peculiares de los usuarios del SBC.
- Inconvenientes
 - Si las tareas habituales no son muy informativas acerca del razonamiento del experto, este método suministra poco conocimiento.
 - Consume mucho tiempo y es a veces inoportuno y fastidioso.

Otras técnicas

- La técnica de las 20 preguntas
 - El IC le presenta un caso inventado a un experto, para resolver el problema el experto solo podrá plantearle 20 preguntas sobre el caso. Sirve para conocer la información relevante que usa el experto en la resolución.
 - Es una variante de la observación de tareas habituales.
- Incidentes críticos
 - Consiste en que el experto describa casos particularmente interesantes o difíciles que se le hayan presentado, además de comentar como lo resolvió.
 - Esta técnica es interesante porque estos casos complejos hacen que el experto comente detalles que en otro momento hubiera ignorado.
 - Una variante es proponer casos críticos imaginarios y ver como el experto lo resuelve.
- Imposición de restricciones
 - Pretende que el experto verbalice ciertas deducciones que en un caso familiar pasaría por alto, imponiéndole limitaciones en el uso de recursos.
 - Se puede restringir el tiempo dado para resolver el problema (Se necesita respuesta rápida) o la información disponible (Faltan datos de entrada).
 - Es útil para educir cierto tipo de conocimiento como lo que pueden ser las prioridades de atención por parte del experto.
- Clasificación de conceptos
 - Se obtienen un conjunto de conceptos del tema a tratar, que cubran ampliamente el dominio. Estos conceptos se transfieren a una ficha y se le pide al experto que los clasifique en una serie de grupos. Describiendo lo que cada grupo tiene en común; luego las fichas de cada grupo se comparan para formar una jerarquía.
 - Esto funciona ya que hay cosas que el experto ya da por obvio, y esto sirve para obtener relaciones entre los conceptos más claros y que le pueden resultar obvias a él.
 - Esta técnica es especialmente aconsejable cuando hay un gran número de conceptos en un dominio, de modo que requieren una estructuración para que sean manejables.
- Cuestionarios

- Consiste en realizar una entrevista estructurada al experto de forma indirecta a través de cuestionarios.
- Frente a las entrevistas, esta técnica tiene la ventaja que es una manera eficiente de acumular información.
- Los cuestionarios pueden ser particularmente apropiados para describir los conceptos, revelar relaciones en el dominio y determinar incertidumbre.