

# Bases de Datos Difusas Sensibles al Contexto

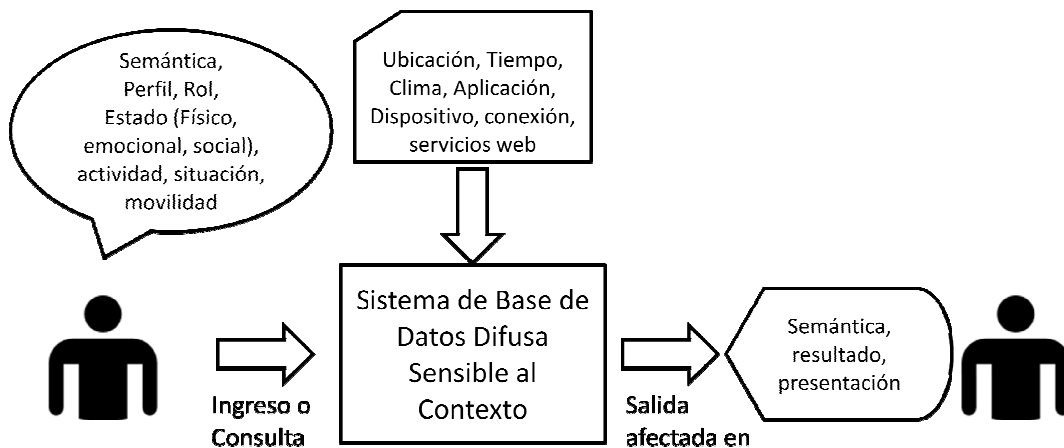
## 3. Modelo Teórico Propuesto

El modelo teórico propuesto es para el desarrollo de sistemas de bases de datos inteligentes basándonos en el modelado orientado a objeto de datos contextuales, la gestión de datos imperfectos mediante bases de datos difusas, reflejando las percepciones del ser humano utilizando la Teoría Computacional de Percepciones, realizando consultas sensibles al contexto que puedan afectar tanto los resultados como la presentación de los mismos. Se quiere que este tipo de sistemas de bases de datos privilegie la utilidad de la información sobre la eficiencia de la consulta.

Frente al paradigma tradicional de operación de un Sistema de Base de Datos donde el usuario proporciona una entrada (ingreso de datos o consulta), se procesa y se obtiene la salida, con la certeza de que siempre que sea la misma entrada tendrá los mismos resultados como salida (paradigma de coincidencia exacta); se propone un enfoque distinto de aplicaciones, donde dada la entrada y el contexto del usuario que proporciona tal entrada, se efectúa el proceso y se pueden obtener resultados diferentes (personalizados) que van a depender no sólo de los datos ingresados sino también del contexto del usuario que interacciona con la base de datos difusa (paradigma flexible de sensibilización al contexto).

### 3.1. Esquema de Interacción

En la figura 1 se presenta el esquema general de la interacción de un usuario con el sistema de base de datos difusa sensible al contexto propuesto. Se puede notar que el sistema va a tomar en cuenta la entrada tradicional (ingreso de datos o consulta), el contexto explícito del usuario (perfil, rol, semántica, estado, actividad, situación y movilidad) y el contexto implícito proporcionado por el ambiente que rodea al usuario que hace la consulta (ubicación, tiempo, clima, aplicación, dispositivo, conexión y servicios web). Es por ello que la salida del sistema puede ser afectada en la semántica de la información presentada, el resultado (tamaño, precisión, orden, clasificación, tipo) y su presentación (formato de salida, tipo de medio).



**Figura 1.** Esquema de interacción de un usuario con el Sistema de Base de Datos Difusa Sensible al Contexto

*Definición.* El esquema de interacción de un usuario con el Sistema de Base de Datos Difusa define los datos contextuales (implícitos y/o explícitos) que al cambiar sus valores pueden afectar la salida (en su semántica, resultado o presentación) de un ingreso o consulta estándar del usuario.

El modelo que se propone es flexible, ya que el diseñador del sistema puede hacer las adaptaciones o transformaciones que considere pertinentes de acuerdo al dominio de una aplicación en particular, es decir, es posible tomar decisiones en cuanto a qué aspectos del contexto (no contemplados aquí) se desean representar o eliminar. Por ejemplo, se puede prescindir del rol o el perfil del usuario sin afectar su propósito. Además, la implementación y puesta en práctica del sistema proporcionará retroalimentación importante para poderlo ajustar, esto es significativo para cualquier aplicación en el ámbito de base de datos donde, luego de evaluar el comportamiento de la aplicación, se puede ajustar el modelo mediante las transformaciones y adaptaciones pertinentes en el diseño del mismo.

Tomaremos como ejemplo a un hospital como dominio de la aplicación. Las características generales relevantes de los hospitales que reflejan el uso de sistemas sensibles al contexto son, entre las que se pueden destacar: la naturaleza distribuida de la información, la necesidad de colaboración, la movilidad (del personal y dispositivos) y la necesidad de tener acceso a información médica precisa en el momento oportuno para la toma de decisiones (Sadri, 2011). Adicionalmente, consideramos que el área de la salud es un ejemplo donde las decisiones están basadas en datos imperfectos y es importante tomar en cuenta las percepciones de diversas personas interesadas involucradas.

En las próximas sub-secciones se presenta el marco teórico de Bases de Datos Difusas, se describe el contexto explícito, el cual es proporcionado por el usuario antes de la consulta o ad-hoc en el momento de la consulta. Luego, el contexto implícito proporcionado por sensores del ambiente, otras aplicaciones o servicios web. Finalmente se describe en qué forma puede ser afectada la salida de la consulta por cambios en el contexto actual del usuario.

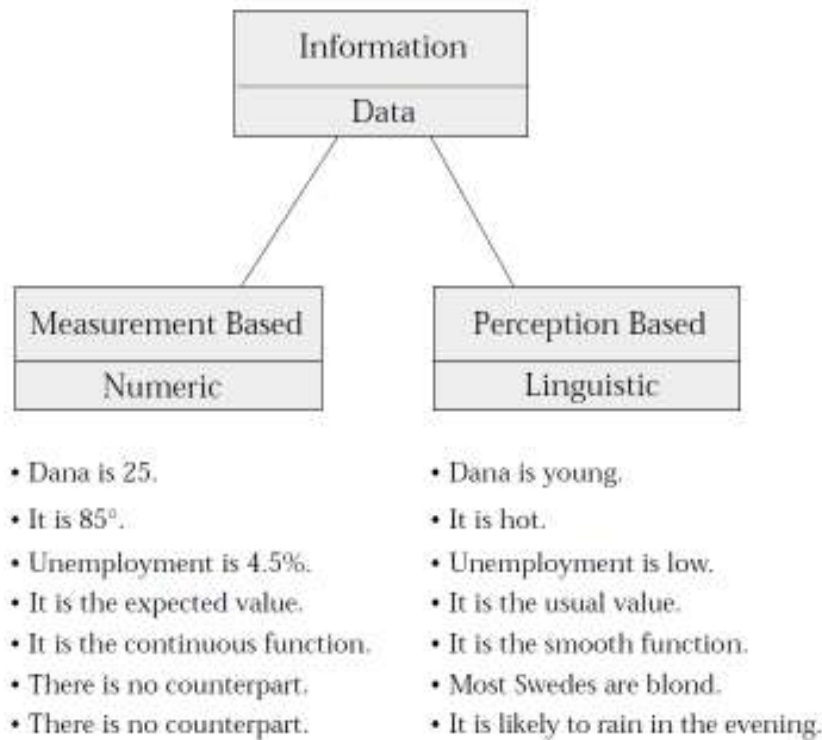
### **3.2. Bases de Datos Difusas**

Como novedad en nuestra propuesta hemos agregado el uso de bases de datos difusas para el desarrollo de aplicaciones sensibles al contexto. Esta incorporación se hace debido a múltiples razones, entre las que destacan:

1. La imperfección de datos contextuales y de los datos del mundo real proporcionados por el usuario
2. La posibilidad de representar percepciones del ser humano por lo que se puede personalizar la semántica utilizada por cada usuario
3. Se pueden obtener consultas flexibles que permiten enfrentar el problema de rigidez de las bases de datos tradicionales afectando el conjunto resultado obtenido por una consulta, minimizando el inconveniente de la respuesta vacía, añadiendo una clasificación a las respuestas (grado de pertenencia) y con la posibilidad de ordenarlos.

Uno de los aspectos innovadores del contexto explícito de nuestro modelo es la posibilidad de definir la semántica del lenguaje utilizado por el usuario de acuerdo a sus percepciones o intereses, lo cual es un avance con respecto a las investigaciones realizadas hasta ahora para el desarrollo de aplicaciones sensibles al contexto.

La Teoría Computacional de Percepciones (CTP) propuesta por Zadeh, proporciona una base teórica para modelar sistemas complejos que gestionen información imperfecta, convirtiéndose en una poderosa herramienta que permite formalizar procesos de naturaleza humana (tal como conversar, razonar y tomar decisiones). Mediante el uso de etiquetas lingüísticas se pueden representar valores imperfectos para ser almacenados en la base de datos o para hacer posible el uso de términos lingüísticos en las consultas que se asemejan más al lenguaje natural utilizado por los seres humanos, es decir, utilizar el enfoque de representar la información en forma lingüística basado en la percepción más que en forma de medidas basado en números (ver figura 2).



**Figura 2.** Datos numéricos versus datos lingüísticos, tomado de Zadeh (2001)

Además, es necesario resaltar la habilidad extraordinaria del cerebro humano para manipular percepciones con respecto a diversos aspectos tales como: distancia, tamaño, peso, color, velocidad, tiempo, dirección, fuerza, verdad, probabilidad y otras características de objetos físicos y mentales; jugando un papel crucial en el reconocimiento de patrones, ejecución de actividades y toma de decisiones.

En la Teoría Computacional de las Percepciones (CTP) propuesta por Zadeh (2001) las etiquetas lingüísticas tienen un rol primordial, expresadas como proposiciones en un lenguaje natural, que son traducidas a un Lenguaje Generalizado de Restricción (GCL) utilizando técnicas basadas en Computación por Palabras (CW). La amplia variedad de restricciones soportadas por GCL (posibilísticas, probabilísticas, grafos difusos, conjunto Pawlak o aleatorio) lo convierten en un lenguaje mucho más expresivo que la lógica de predicados.

Las percepciones según Zadeh (2001) son *f-granular*, lo que significa que los límites de las clases percibidas son imprecisos o borrosos. Los valores de los atributos son granulados, con

un gránulo representado por un grupo de valores (puntos, objetos), agrupados por su semejanza, proximidad o función. Por ejemplo, los gránulos de una variable lingüística como la edad podrían ser las etiquetas lingüísticas: muy joven, joven, promedio, viejo, muy viejo. La importancia de CTP deriva del hecho de que en gran parte la toma de decisiones y el razonamiento común está basado en percepciones humanas.

Mediante la utilización de conjuntos difusos se puede identificar aspectos del mundo real del usuario y con la especificación de una adecuada función de pertenencia (la cual es subjetiva y dependiente del contexto) u otro dominio difuso que permitan establecer matrices de semejanza, se puede definir la semántica de conceptos abstractos de acuerdo a la percepción individual de cada usuario, logrando personalizar el lenguaje utilizado por cada usuario.

Otro punto importante afirmado por los autores Berzal, Marín, Pons y Vila (2007) es que muchas formas de imperfección involucran la noción de “gradualidad”, ya que en diversas situaciones los conceptos utilizados por el ser humano para describir una situación no corresponden a una realidad simple y/o precisa, sino a una gradual. Por ejemplo, el término *joven* es gradual ya que una población dada no se puede dividir sólo en dos conjuntos: los jóvenes y los que no lo son porque existe un tercer conjunto de personas en las fronteras de estos dos que pueden tener cierto grado de joven, es decir, algunas personas indudablemente son jóvenes, otras no lo son y hay un tercer grupo que está entre estos dos, donde los límites de los tres conjuntos no están precisamente definidos, de allí el término de borroso o difuso para estos tipos de conjuntos.

Esta noción de gradualidad nos permitirá abordar el problema de rigidez del paradigma de consulta exacta, al clasificar las respuestas y establecer un orden de prioridad (en el caso de muchas respuestas), además de poder flexibilizar las condiciones de la consulta y ofrecer respuestas que no cumplan estrictamente las condiciones requeridas (en el caso de consultas sin respuestas) sino que se aproximen a las necesidades del usuario con un cierto grado de pertenencia.

Aquí es vital para la propuesta desarrollada en esta investigación tomar en cuenta la noción de contexto, ¿Qué persona puede establecer los límites (aunque sean difusos) para determinar por ejemplo que una persona es indudablemente, más o menos *joven* o definitivamente no lo es? El término *joven* puede estar influenciado por aspectos contextuales tales como la ubicación de la persona (en qué país está, por ejemplo España o en Venezuela), por los objetivos o intereses del usuario (si se quiere un gimnasta o un investigador) o por otros factores como la percepción de la persona que está haciendo la definición del término, entre otros elementos que pueden ser parte del contexto y afectar la semántica del concepto.

### **3.2. Contexto explícito**

El usuario de un sistema de base de datos debe estar siempre identificado debido a los niveles de seguridad que garantizan que los datos puedan ser accedidos sólo por aquellos que tengan privilegio para hacerlo. Es por ello que un sistema de base de datos siempre tiene una primera información del contexto que es la identificación del usuario que interacciona con el mismo.

A continuación se presenta un análisis de los aspectos considerados del contexto explícito que pueden afectar la salida de una consulta en nuestra propuesta de Sistemas de base de datos difusa sensibles al contexto.

### **3.2.1. Semántica**

Al conectarse un usuario al sistema de base de datos puede definir las semánticas de las etiquetas que desea utilizar de acuerdo a sus necesidades o preferencias, dejando abierta la posibilidad de que un usuario por defecto (diseñador del sistema o usuario experto) haga estas definiciones.

De esta forma tanto la semántica como los resultados de las consultas pueden ser afectados (relevancia, orden o clasificación de los mismos). También, se pueden obtener filas con un grado de pertenencia con respecto a los términos lingüísticos involucrados en la consulta, esto es lo que denominamos consulta flexible o difusa.

Una de las forma de gestionar datos imperfectos a través de un DBMS es colocando capas lógicas de software por encima de éste, que haga las veces de interfaz con el usuario, traduciendo posteriormente a sentencias del DBMS (normalmente SQL) y luego traducir de nuevo los resultados a un ambiente flexible, esto supone costos adicionales que se manifiestan en un impacto en el rendimiento de estos sistemas.

Nuestra propuesta es utilizar las capacidades Objeto Relacional de gestores de base de datos para, mediante la extensión de los tipos nativos de datos, gestionar los dominios difusos propuestos por Marín et al. (2003). Además, con el uso del lenguaje nativo de procedimientos y funciones de SQL lograr un bajo impacto en el rendimiento del sistema; así como afectar, en lo mínimo posible, el tipo de sentencias estándar utilizado por el usuario en su interacción con el sistema, mediante aplicaciones amigables y donde la sensibilización al contexto sea lo más transparente posible. Una primera aproximación del diseño e implementación de esta semántica adaptable al contexto del usuario utilizando dominios difusos con el ORDBMS Oracle se muestra en Cadenas, Marín y Vila (2011).

### **3.2.2. Perfil**

En nuestra propuesta el usuario al conectarse al sistema puede especificar un perfil explícitamente, mediante un cuestionario donde especifique datos personales (tales como fecha de nacimiento, intereses, hábitos, área de trabajo, entre otros); o el perfil puede ser inferido (evaluando las interacciones del usuario con las aplicaciones, por ejemplo desde qué ubicación o aplicación hace la interacción con la Base de Datos). Este perfil de usuario permitirá dar relevancia, filtrar y personalizar las respuestas a los usuarios.

Por ejemplo, el médico de un hospital de acuerdo a su experiencia y su área de especialización puede requerir de información relevante según su parecer o con cierto grado de precisión. En un hospital, dependiendo de la ubicación donde se está conectando una persona puede asociársele un perfil de personas que se conecten desde la misma ubicación, a pesar de que el usuario no especifique un perfil explícitamente.

### **3.2.3. Rol**

El rol en un sistema de base de datos tiene un significado muy concreto. Es usual asociar un grupo de usuarios a un rol para conceder permisos a acceder cierta parte de la base de datos o funcionalidades del DBMS, de esta forma el administrador de la base de datos no necesita conceder los permisos en forma individual, sino por grupo de personas.

En nuestra propuesta podemos asociar igualmente un rol para que el contexto funcione para un grupo de personas. Por ejemplo, en el caso de un hospital donde podemos tener cientos de usuarios, se pueden agrupar los mismos por médicos referentes, médicos evaluadores, terapeutas físicos (fisioterapeutas), enfermeras, administradores del centro y familiares de pacientes, entre otros.

A cada rol le podemos asociar la semántica de los conceptos utilizados en variables lingüísticas como el peso, la edad, la talla o tipo de diagnóstico; afectando los resultados por relevancia, orden y precisión, entre otros.

#### **3.2.4. Estado (Físico, emocional, social)**

El estado actual del usuario ya sea en lo físico, emocional o social, nos va a permitir afectar los resultados de las consultas y su forma de presentación, así como en la semántica de los conceptos utilizados.

La palabra “estado” según el Diccionario de la Real Academia Española (DRAE, 2013) proviene del latín *status* y significa “Situación en que se encuentra alguien o algo, y en especial cada uno de sus sucesivos modos de ser o estar”, cuando hablamos de estado físico se refiere a “situación en que se encuentra alguien respecto a su organismo físico”

El paciente de un hospital tiene una evaluación de su estado físico y esto determina cierto nivel de gravedad que puede dar mayor o menor relevancia a los resultados de una consulta acerca del paciente, o en todo caso requerir mayor o menor precisión en las respuestas acerca de sus evaluaciones para que el médico pueda dar un diagnóstico sin afectar la salud del paciente. También si el paciente tiene algún impedimento físico las respuestas pueden ser desplegadas por diferentes medios (si el paciente tiene problemas de visión las respuestas pueden ser auditivas en vez de textuales).

Más aún el estado físico del paciente puede variar continuamente. Nuestro modelo prevé la interacción con sensores u otras aplicaciones, ya que este tipo de contexto es dinámico (Oliver y Flores-Mangas, 2006). Por ejemplo, tecnologías de monitoreo de la frecuencia cardíaca pueden ofrecer información relacionada con la salud de la persona (Ahtinen, Mantjarvi y Hakkila, 2008). En este caso nuestro modelo puede apoyarse en técnicas como bases de datos temporales, almacenes de datos, minería de datos y flujo de datos (*data stream*) combinado con la gestión de datos difusos (por su imperfección), entre otras, para poder gestionar gran cantidad de datos y obtener el conocimiento relevante para los médicos.

En cuanto al estado emocional o de ánimo, según el DRAE (2013) significa “Disposición en que se encuentra alguien, causada por la alegría, la tristeza, el abatimiento, etc”. La primera aproximación en nuestro modelo es que el usuario (paciente en el caso del hospital) diga explícitamente su estado de ánimo (ira, alegría, tristeza, miedo, disgusto y sorpresa, entre otros).

En áreas como la de Interacción Humano-Máquina se han utilizado diferentes técnicas para reconocer las emociones humanas, donde algunas propuestas se han concentrado en las expresiones faciales y el lenguaje, otras en parámetros fisiológicos como la frecuencia cardíaca y la actividad electro-dérmica (Sardi, 2011); también se ha tratado de prever el comportamiento de una persona utilizando vídeos y sensores, mediante técnicas de percepción inteligente (Oliver y Pentland, 2000). Además, dentro de Inteligencia Ambiental la computación afectiva

estudia como habilitar a las computadoras para reconocer las emociones humanas y para actuar en consecuencia (Sadri, 2011). Aunque no está previsto utilizar estas técnicas dentro de nuestro modelo, puede hacerse estudios posteriores para analizar la factibilidad de incorporar alguna de ellas.

El estado emocional del paciente puede influenciar en la relevancia de las respuestas que el médico requiere de un sistema para poder hacer un diagnóstico más ajustado a la realidad, en especial cuando el paciente padece de problemas psiquiátricos tales como esquizofrenia, autismo, bipolaridad o profunda depresión (Hyman, 2003).

En cuanto al estado social, es importante conocer las personas que rodean al usuario (está sola, acompañada con parientes o amigos) o si, por ejemplo, está inmerso en el tráfico de una ciudad. En el caso de pacientes de un hospital, el mismo puede estar en diversos estados sociales e, igual que en el caso anterior, este estado puede influenciar en la relevancia de los resultados de las consultas y su forma de presentación. El médico también puede estar en un particular estado social (por ejemplo, acompañado por otros médicos de diversa especialización o sólo), tanto para hacer diagnósticos o intercambiar información acerca de tratamientos.

Hay aplicaciones hoy en día que toman en cuenta el aspecto social no sólo como identificación del contexto del usuario sino para hacer búsquedas de personas que puedan tener respuestas a interrogantes (Martinenghi y Torlone, 2009), esto puede ser aplicado a médicos que quieren buscar si hay otros médicos que han efectuado tratamientos y sus resultados para un diagnóstico particular. Aquí toma especial interés el uso de herramientas de comunicación social de la web 2.0.

### **3.2.5. Actividad**

Según el DRAE (2013) actividad se define como “Conjunto de operaciones o tareas propias de una persona o entidad”. Con este aspecto se quiere identificar qué tipo de actividad está haciendo el usuario cuando efectúa la consulta (por ejemplo caminando, haciendo ejercicios, en bicicleta, manejando, efectuando una operación quirúrgica, en cama, entre otros).

Algunos teléfonos inteligentes modernos están equipados con una amplia variedad de sensores, incluyendo sistemas de posicionamiento global (GPS) y acelerómetros, éstos se pueden utilizar para detectar el número de pasos y la cadencia al caminar (numero de pasos por minuto), e incluso para clasificar la actividad (distinguir si está caminando o en bicicleta, por ejemplo), lo cual requiere identificar localización y orientación (Schneider, 2013).

En el ámbito de la salud, esto puede ser usado para ingresar datos en una aplicación de diagnóstico para laboratorios en marcha para un estudio completo en el caso de disfunciones neuromotoras (Parálisis Cerebral, Espina Bífida, Mielodisplasia, malformaciones congénitas, entre otras). La actividad detectada del paciente unido a definiciones semánticas de tipos de marcha, puede ayudar a realizar consultas difusas que involucran preferencias de usuarios como una herramienta poderosa para la toma de decisiones de los médicos (Aguilera y Rodríguez, 2011).

Por otro lado es importante identificar la actividad de un médico: si está realizando una operación quirúrgica, está en su despacho haciendo consultas a pacientes o está de descanso. Un paciente puede estar en un estado crítico, fuera de peligro o en estados intermedios que pueden ser definidos a través de etiquetas lingüísticas por los médicos o expertos. Esto puede ayudar a

obtener respuestas relevantes o con mayor o menor precisión a consultas difusas. Un médico que está haciendo una operación quirúrgica quisiera que su interacción con el sistema sea en forma auditiva.

### **3.2.6. Situación**

De acuerdo al DRAE (2013), la situación es un “conjunto de factores o circunstancias que afectan a alguien o algo en un determinado momento”. Es por ello que tomamos la situación como las circunstancias que rodean al usuario cuando efectúa una consulta. Es muy importante este aspecto en un sistema sensible al contexto, las personas utilizan el conocimiento de las circunstancias actuales para estructurar actividades, interactuar con el mundo que los rodea, organizar información y adaptarse a las condiciones. Por ejemplo, cuando se mantiene una conversación en un lugar ruidoso, se habla más alto para que la otra persona pueda oír; pero cuando se está en una conferencia, susurramos para no molestar a los demás participantes.

Un médico en una unidad de asistencia de emergencia que debe atender a un paciente que llegó por un accidente de tránsito, le sería útil conocer información acerca del historial clínico del paciente (Jiang y col., 2010). Un paciente puede estar en un hospital en una unidad de cuidado intensivos o en una habitación normal o en su casa, dependiendo de su situación puede ser afectada la relevancia de los resultados de las consultas.

### **3.2.7. Movilidad**

Por movilidad se entiende la capacidad que tiene la persona de desplazarse por sus propios medios. Hay personas que pueden tener limitada la capacidad de moverse sin ayuda externa, este grupo se puede componer tanto de aquellos que tienen una discapacidad relacionada con la movilidad (paraplejía, tetraplejía, problemas óseos, entre otros) como quienes tienen dificultades por otros motivos (ancianos, embarazadas, personas con niños pequeños, entre otros). También aplica para personas que deben trasladarse a pie, o tienen posibilidad de trasladarse en coche o a través de algún servicio de transporte (taxi, autobús, tren, avión).

Como dato contextual en nuestro ejemplo de hospital puede servir para conocer si debido a la ubicación de un médico específico puede trasladarse a un hospital o deba hacerse por ejemplo una video conferencia. En el caso de pacientes (ya como dato imperfecto del mundo real) es importante este aspecto para saber qué tipo de tratamientos pueden ser aplicados y también para los diagnósticos. La clasificación de la movilidad puede hacerse por etiquetas lingüísticas con dominios categóricos por lo que se necesita una matriz de semejanza para establecer comparaciones entre etiquetas y que sea aplicable en las consultas difusas.

## **3.3. Contexto Implícito**

En este apartado el sistema de base de datos debe ser consciente de aspectos dinámicos del ambiente que puedan afectar la interacción son el sistema de base de datos sensible al contexto.

De acuerdo a Levandoski, Khalefa y Mokbel (2010) en la búsqueda de restaurantes, un usuario puede desear tomar en cuenta el tiempo de viaje (utilizando el servicio web *Bing Maps*) o reseñas de restaurantes (utilizando el servicio web *Yelp*). Un problema primordial es cómo integrar este tipo de datos contextuales dinámicos en el procesamiento de consultas, por lo que estos autores plantean dos dificultades principales: (1) los atributos contextuales son más costosos de derivar con respecto a los datos estáticos almacenados localmente en una base de



datos (por ejemplo, extraerlos de un tercero en una red) y (2) Los datos contextuales pueden ser imperfectos (por ejemplo, los precios del restaurante son informados como un rango y no en forma precisa).

En nuestro modelo abordamos el primer problema con la incorporación de técnicas de bases de datos temporales, bodegas de datos, flujo de datos y minería de datos. Mientras que el segundo es enfrentado mediante la utilización de bases de datos difusas. Por otro lado para la adquisición de datos implícitos se piensa en un primer paso aprovechar la cantidad de datos contextuales y servicios web que puedan ser accedidos a través de dispositivos móviles tales como teléfonos inteligentes y tabletas digitales, por lo cual se hace necesario interfaces de comunicación con estos tipos de dispositivos.

### ***3.3.1. Ubicación***

La ubicación puede describir no sólo el lugar físico donde está la persona (ciudad, país, coordenadas) proporcionadas por sistemas tales como GPS; además, puede proporcionar información espacial relacionada tal como la velocidad u orientación, incluyendo información de la zona en que se encuentra el usuario. En general los datos obtenidos por este tipo de contexto pueden ser proporcionados (además de sensores) por otras aplicaciones o servicios web de terceros.

En cuanto a nuestro ejemplo médico la ubicación puede servir para identificar desde donde se están conectando los usuarios que consultan (puede ser por direcciones IP) para asociarlos a algún perfil.

### ***3.3.2. Tiempo***

En la sección 3.2 se ha mencionado que hay casos en que podemos obtener datos a través de sensores que están asociados al tiempo (fecha y hora) que se adquieren. En este sentido, nuestro modelo debe usar técnicas del enfoque de bases de datos temporales para almacenar y gestionar este tipo de datos que se pueden generar en gran cantidad.

En el caso de la aplicación médica es un hospital es importante este contexto para cuando las enfermeras o personal médico haga consultas acerca de los tratamientos que deben ser aplicados o el calendario de intervenciones quirúrgicas.

### ***3.3.3. Clima***

El clima al igual que la ubicación puede ser proporcionado por servicios web y es común que estos datos estén disponibles en un dispositivo móvil inteligente. El clima puede tener influencia en el tipo de tratamiento que se le dé a un paciente (que puede efectuar actividades al aire libre o no) en nuestro dominio de aplicación de la salud, también puede influir en una consulta de los médicos disponibles de acuerdo a su ubicación y movilidad.

### ***3.3.4. Aplicación***

Al igual que un usuario, otra aplicación puede conectarse con nuestro sistema para hacer consultas. En este caso es importante conocer el tipo de aplicación que está haciendo la consulta y de esta forma puede asociarse a un rol en particular. Por ejemplo puede ser una aplicación que ingresa datos de exámenes específicos como el físico articular, donde la interacción es efectuada por fisioterapeutas que tienen una semántica definida de acuerdo a su rol.

### **3.3.5. Tipo de dispositivo**

Por tipo dispositivo entendemos si el usuario se conecta desde un dispositivo fijo (computador de un despacho), portátil o móvil (teléfono inteligente, tableta). Este contexto permitirá presentar los resultados en forma diferente (resumidos, a través de texto o multimedia). Aplicable a cualquier tipo de consulta en cualquier dominio de aplicación.

### **3.3.6. Tipo de conexión a la red**

El tipo de conexión a la red define la velocidad a la que se está conectando a través de etiquetas lingüísticas tales como lento, medio o rápido (éstas deben ser definidas por el usuario o diseñador del sistema). Este es un factor que también determina la presentación de resultados en diversos medios, por ejemplo si la conexión es lenta es preferible sólo texto que multimedia. Aplicable de nuevo a cualquier tipo de consulta en cualquier dominio de aplicación.

### **3.3.7. Servicios Web**

En cuanto a la obtención de información de servicios web médicos, es importante la información de ontologías del área médica, tales como las de enfermedades como el cáncer, o sitios web de bibliotecas digitales temáticas (*Digital Library*) consideradas confiables por los profesionales del hospital.

Para trabajos futuros se puede pensar en la incorporación en el modelo de técnicas aplicadas en diversos enfoques tales como “Internet de las cosas”, “Web<sup>2</sup>”, “Web de sensores” o “Inteligencia Ambiental” lo que ampliaría la aplicación de nuestra propuesta.

## **3.4. Salida sensible al contexto**

El contexto puede afectar los resultados producidos por una consulta, en estos casos, una misma consulta realizada en diferentes contextos puede dar resultados diferentes aún si los datos contenidos en la base de datos no han cambiado. A continuación, se describe los aspectos de la salida de una consulta que consideramos pueden ser afectados en el esquema de interacción de un usuario con una base de datos debido a cambios en el contexto.

### **3.4.1. Semántica**

Como ya se explicó en el apartado anterior, la utilización de la Teoría Computacional de Percepciones y Bases de Datos Difusas permite que nuestra propuesta pueda representar conceptos del mundo real de acuerdo a las percepciones de cada usuario. Así, a pesar de que los datos almacenados sean efectuados por un usuario con una semántica en particular, nuestro sistema es capaz de transformar lingüísticamente esos conceptos de acuerdo al usuario que hace la consulta. Esto es una novedad con respecto a cualquier propuesta de sistema de base de datos sensible al contexto.

En el artículo de Cadenas, Marín y Vila (2011) se muestra un ejemplo de cómo el usuario puede definir semánticas de etiquetas lingüísticas, almacenando objetos tipo trapecio en la base de datos difusa de acuerdo a estas definiciones. Cuando otro usuario, que tiene definiciones distintas de las etiquetas relacionadas con el concepto, se conecta al sistema y efectúa una consulta las respuestas pueden ser mostradas de acuerdo a la semántica de este último. Es así como el usuario (contexto) influye en la semántica de conceptos lingüísticos almacenada en la base de datos y en los resultados sensibles al contexto

Este ejemplo lo hemos generalizado para que funcione no sólo con trapezios (etiquetas lingüísticas donde el dominio básico referente es continuo) sino para otros dominios difusos (etiquetas lingüísticas sin representación semántica asociada, etiquetas lingüísticas donde el dominio básico referente es discreto o conjuntos difusos), como lo explicaremos en próximas secciones.

En el dominio de aplicación de la salud, en un hospital donde hay diversos usuarios (médicos referentes, médicos evaluadores, terapeutas físicos, pacientes) cada uno puede tener definiciones de sus propias etiquetas sobre variables lingüísticas (peso, talla, edad, rangos articulares de una persona, entre otros) y al efectuar consultas difusas los resultados se pueden presentar con semántica lingüística transformada de acuerdo al usuario que hace la consulta.

### **3.4.2 Resultado**

Uno de los objetivos del sistema de base de datos sensible al contexto es ajustar o personalizar el resultado de acuerdo al contexto para flexibilizar las respuestas que se obtienen de la consulta, abordando los problemas de rigidez de la consulta que se producen con el paradigma de coincidencia exacta. Esto incluye diversas características del resultado que hemos identificado en nuestro modelo, a saber: tamaño, precisión, orden, clasificación y tipo, a continuación se describe el significado de cada uno de ellos.

#### **3.4.2.1. Tamaño**

Por tamaño entendemos el número de filas en la salida del resultado. Así, dependiendo del tipo de dispositivo con que se conecte el usuario (computador de escritorio o dispositivo móvil) se puede presentar un número diferente de filas (más o menos respectivamente). En cuanto al número de columnas en nuestra propuesta no va a ser afectado porque nos proponemos *utilizar* el lenguaje SQL estándar de un ORDBMS y en la sentencia SELECT de la consulta ya están identificadas las columnas que solicita el usuario como salida.

Otro aspecto del contexto que puede influenciar el tamaño de la salida es la semántica de los conceptos utilizados. Por ejemplo, cada usuario puede definir etiquetas para la altura de las personas y al efectuar consultas difusas donde se involucre en la condición dichas etiquetas se obtendrán más o menos filas en las respuestas con diferentes grados de pertenencia.

En nuestro caso médico hay muchos ejemplos de utilización de etiquetas lingüísticas. Por ejemplo el tono muscular (capacidad de contracción de un músculo), se pueden representar con las etiquetas *atonía* (*falta de tono*), *hipotonía*, *normotonía*, *hipertonía* (*exceso de tono con espasticidad*). Si además se definen etiquetas como *infante*, *niño*, *joven*, *adulto*, *viejo*, sobre la variable lingüística edad; *ligero*, *liviano*, *medio*, *pesado*, *obeso* sobre el “peso”; *bajo*, *mediano* y *alto* sobre la “altura”. Luego se pueden hacer consultas difusas tales como:

- Pacientes *viejos* y *obesos* con *atonía* en extensores rodilla izquierda
- Pacientes *jóvenes* con *hipotonía* en los flexores rodilla derecha
- Pacientes *viejos* que tienen *atonía* en la flexión de la cadera derecha
- Pacientes *ligeros* que tienen *normotonía* en los plantares izquierdos

Donde la cantidad de filas del resultado va a ser afectado por la semántica. También en las consultas difusas el usuario puede colocar un alfa-corte para restringir el número de

resultados y obtener como filas sólo las que cumplan con un grado de pertenencia a la consulta mayor o igual a cierto número entre 0 y 1.

El perfil del usuario es otro aspecto del contexto que puede afectar el tamaño del resultado de una consulta sensible al contexto. Esto es porque en el perfil se pueden definir diferentes parámetros contextuales (como intereses, hábitos, país donde reside, entre otros), que al incorporarlos a la consulta funcionan como restricciones adicionales que van a tener un efecto sobre el número de filas que cumplen la condición. En el caso médico las etiquetas lingüísticas de rangos articulares normales (límites de grados en que el paciente logra realizar el movimiento con cualquier articulación como el pie, rodilla o cadera) de un médico pueden ser diferentes a las de otro médico o especialista.

En cuanto al rol, como mencionamos en el apartado anterior, puede servir para identificar conjuntos de usuarios que comparten intereses, semántica de conceptos o perfiles que van a influenciar en el tamaño del resultado de una consulta consciente del contexto. En el área de la salud, la fuerza muscular (capacidad del músculo para vencer una fuerza que se opone al movimiento) puede ser representada con etiquetas como *muy poca*, *poca*, *moderada*, *buena* y *mucha*, cuyo significado puede variar para especialistas como los fisioterapeutas con respecto a las definidas por los médicos cirujanos.

Por otro lado, la ubicación desde donde se hace la consulta, puede ayudar a ser más selectivo con las respuestas a la consulta. En el caso médico podemos identificar un usuario con un perfil de acuerdo al sitio de donde hace la consulta (a pesar de que no tenga un perfil explícitamente definido), es decir, ciertas direcciones de IP del hospital pertenecen a los computadores de los laboratorios donde los fisioterapeutas efectúan los exámenes físicos, mientras otras direcciones IP son de los computadores ubicados en los consultorios de los médicos que hacen el diagnóstico.

Adicionalmente al tamaño puede ser combinado con otros aspectos del resultado como la relevancia para presentar en forma priorizada el mismo, lo cual será explicado posteriormente.

### **3.4.2.2 Precisión**

La precisión es un aspecto del resultado que nos permitirá flexibilizar en mayor o menor medida las condiciones de la consulta, permitiendo obtener más o menos respuestas respectivamente, por lo tanto incidiendo en el tamaño del resultado.

La situación de un usuario es una característica del contexto que va a influenciar la precisión. Por ejemplo un usuario que consulta acerca de noticias, le interesará las relacionadas al tráfico cuando está manejando o el clima cuando está caminando o en bicicleta. En el caso del hospital los resultados de una consulta acerca de un diagnóstico de un paciente deben ser más precisos cuando el paciente está en estado grave (por ejemplo, en terapia intensiva) que cuando está en su casa.

El uso de etiquetas lingüísticas en la condición de la consulta (condiciones difusas) y el de atributos con dominio difusos en la base de datos, permite obtener respuestas con un grado de pertenencia, donde éste indica el grado de cumplimiento de las respuestas con respecto a las condiciones difusas. Cuando el usuario desee un alto grado de precisión en las respuestas debe exigir que este grado sea 1 o un alfa corte alto. En casos más relajados, donde incluso se puede

tener el problema de que ninguna respuesta cumpla con todos los requisitos el grado de pertenencia servirá como medida de precisión para obtener respuestas que se acerquen a las necesidades del usuario, estableciendo alfa-cortes más flexibles.

#### **3.4.2.3. Orden**

El orden en que se presentan los resultados es importante cuando se tiene muchas respuestas y se quiere priorizar o jerarquizar las ajustadas a las necesidades de los usuarios. Así aspectos como la localización y la movilidad del usuario (si va a pie) pueden influir en el orden de los resultados en consultas tales como encontrar los restaurantes cercanos.

En el caso de los pacientes de un hospital, la movilidad va a influenciar en el tratamiento que se dé como respuesta a un médico, por ejemplo un paciente con movilidad reducida se va a ordenar la respuesta por tratamientos que requieran menor movilidad.

Aquí también puede jugar un papel importante el uso de etiquetas lingüísticas para generar consultas difusas. Dado que al colocar condiciones difusas se va a obtener respuestas con un grado de pertenencia a la consulta (implementando el concepto de gradualidad explicado anteriormente), donde mientras mayor sea el grado más se acerca al cumplimiento de las necesidades del usuario, de allí que ordenar por este grado nos dará las respuestas ordenadas de acuerdo a la relevancia según necesidades del usuario.

#### **3.4.2.4. Clasificación**

En un dominio de aplicación donde se pueden obtener muchas respuestas a las consultas, es importante establecer algún criterio para clasificar las mismas, lo que se desea es poder de alguna forma jerarquizar las respuestas de acuerdo a las necesidades del usuario, logrando la personalización de las mismas. De nuevo, en nuestra propuesta, al tener almacenados datos imperfectos y utilizar etiquetas lingüísticas para atributos de datos precisos podemos obtener consultas difusas donde el grado de pertenencia en la respuesta nos permite clasificar las respuestas de acuerdo al grado de cumplimiento a las condiciones de la consulta (necesidades del usuario).

En nuestro ejemplo médico, en el caso de pacientes donde el peso es descrito por etiquetas lingüísticas (*ligero, liviano, medio, pesado, obeso*), además el rango articular de la cadera (por ejemplo) también es descrito por etiquetas como *bajo, normal* y *alto*; se puede combinar esto en una consulta difusa tal como “personas obesas que tengan la flexión de la cadera derecha dentro de los rangos normales”. Obteniendo respuestas con un grado de pertenencia que clasifican las mismas de acuerdo a las necesidades del usuario.

#### **3.4.2.5. Tipo**

El tipo de resultado se refiere a que la respuesta puede ser condicionada a un contexto dependiendo del dominio de la aplicación. Por ejemplo consultas pueden dar resultados diferentes debido al interés o disposición de tiempo del usuario dependiendo del tipo de día en que se hace (laboral, fin de semana o día feriado), o cambios del estado social (solo, acompañados con amigos, acompañado con familia) puede afectar los resultados de búsqueda del tipo de sitio o evento de preferencia (por ejemplo un tipo de restaurant, cine o teatro).

En nuestro caso médico el tono muscular puede depender de la edad de la persona, así podemos hacer consultas tales como: Pacientes *viejos* que tienen *atonía* en la flexión de la

cadera derecha, donde viejo y atonía son etiquetas lingüísticas definidas por el usuario según sus necesidades para las variables lingüísticas edad y flexión de la cadera respectivamente.

### 3.4.3. Presentación

En cuanto a la presentación de los resultados es claro la influencia del contexto en dos aspectos principales:

1. El formato de salida, el cual básicamente entendemos si se va a presentar en un formato estándar de computador de escritorio o para dispositivos móviles (un formato más resumido).
2. El tipo de medio, el cual nos indicará si la salida se la podemos presentar al usuario en texto, multimedia (texto con imágenes, vídeo y/o sonido) o un solo medio (sonido por ejemplo).

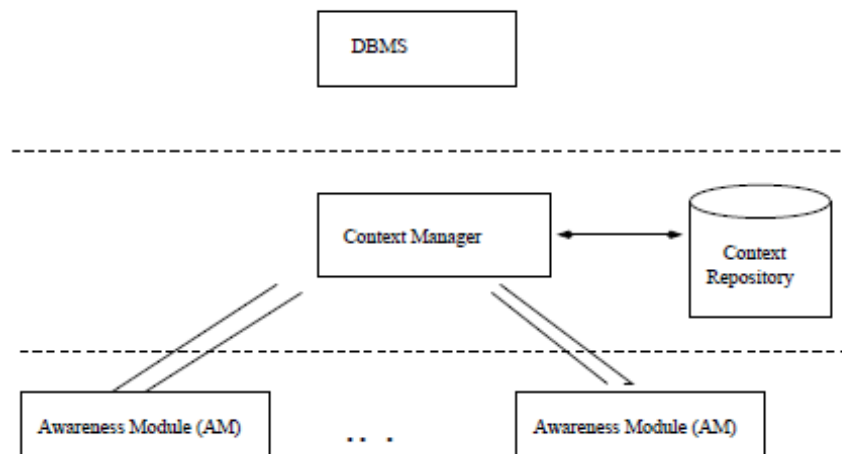
Dependiendo de la actividad, por ejemplo si el usuario está manejando o va caminando, la respuesta puede ser desplegada en sonido o multimedia respectivamente. Dependiendo de la calidad de la conexión a la red o el tipo de dispositivo también nos dará la posibilidad de usar tipos de medio distintos para presentar los resultados.

En nuestro dominio de aplicación de la salud es claro que el resultado se puede presentar diferente cuando el médico efectúa la consulta desde su despacho atendiendo a pacientes o si hace la consulta desde su dispositivo móvil cuando hace una ronda por el hospital.

### 3.5. Sistema de Base de Datos Difusa Sensible al Contexto

Una vez conocido el esquema de interacción de un usuario donde los ingresos y consultas van a ser conscientes al contexto pudiendo afectar la salida, vamos a plantear cómo sería el modelo teórico del funcionamiento interno del sistema sensible al contexto. Posteriormente explicaremos con una prueba de concepto la factibilidad de su funcionamiento.

En la Figura 3 se muestra el esquema general propuesto por Pitoura, Stefanidis y Zaslavsky (2004) para integrar el contexto en un Sistema Gestor de Base de Datos (DBMS).



**Figura 3.** Conectando el Contexto y las Bases De Datos

Los componentes que se muestran en la figura son los siguientes:

- Los “Módulos de Sensibilización” (*Awareness Modules*) se comunican con las fuentes que producen los datos (por ejemplo, sensores) y propagan las actualizaciones al Gestor de Contexto.
- El “Gestor de Contexto” (*Context Manager*) es responsable por administrar (modelar, almacenar, actualizar) la información relacionada al contexto.
- El “Repositorio de Contexto” (*Context Repository*) es el módulo donde el contexto es almacenado.

Hay dos formas de integrar el contexto en un DBMS: a) El gestor de contexto puede ser parte del DBMS o b) El gestor de contexto puede ser visto como un mediador en una capa intermedia.

Nosotros ideamos el Gestor como una capa intermedia, pero con la ventaja de utilizar las capacidades de orientación a objeto y el lenguaje nativo estándar procedimental del ORDBMS, con dos propósitos principales: 1) Utilizar sentencias estándar, tratando de ser lo menos intrusivo posible en las sentencias utilizadas por los usuarios y 2) Afectar en menor medida (que las propuestas de capa intermedia hechas hasta ahora) el rendimiento del ORDBMS ofreciendo nuevas funcionalidades que le dan flexibilidad a las consultas de los usuarios.

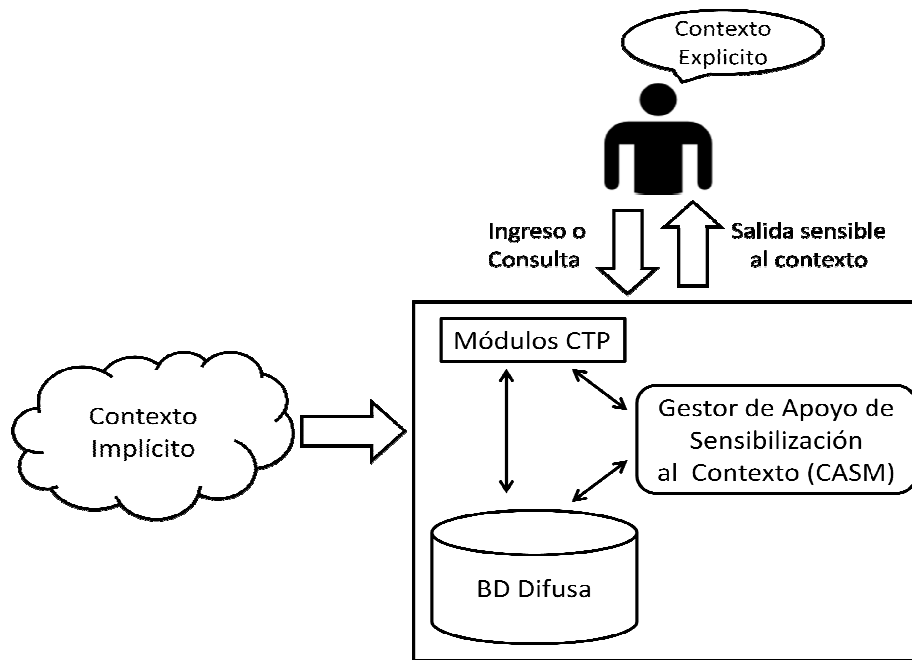
En nuestra propuesta la innovación que estamos introduciendo es la utilización de una Base de Datos Difusas en la capa del DBMS. Así podremos almacenar datos imperfectos mediante la tecnología de orientación a objetos que luego podemos comparar a través de la utilización de métodos, funciones y procedimientos, utilizando el lenguaje nativo estándar del ORDBMS, evitando colocar mediadores en capas intermedias en lenguajes no nativos que puedan impactar significativamente el rendimiento del sistema en general.

Así, diseñamos dos Módulos de sensibilización al contexto los cuales son: el módulo de almacenamiento (CAS) y el módulo de recuperación (CAR), ambos cumplen con los postulados de la Teoría Computacional de Percepciones para permitir ingresar y consultar datos a través de etiquetas lingüísticas, que pertenezcan a algún dominio difuso y sean conscientes del contexto.

El modulo CAS permite la inserción de datos de forma amigable a un usuario, mediante la implementación de métodos constructores para los dominios difusos definidos en la base de datos y la extensión de tipos de datos nativos de datos en un ORDBMS. Mientras que el CAR permite la recuperación de datos con sensibilidad al contexto, mediante la utilización de métodos, funciones y procedimientos implementados en el lenguaje nativo del ORDBMS.

El Gestor de contexto lo denominamos Gestor de Apoyo a la Sensibilización al Contexto (*Context Awareness Support Manager*). Este módulo es el encargado de gestionar el Catalogo Contextual, el histórico de las acciones de los usuarios y reglas que permitan inferir comportamiento de acuerdo al contexto. Los datos contextuales pueden ser imperfectos por lo cual se utilizará para su gestión un modelo de datos orientado a objeto difuso. Este gestor constituye la meta base de conocimiento para que los Módulos CTP puedan actuar como agentes inteligentes y realizar la sensibilización al contexto de los datos almacenados o provenientes de la base de datos difusa.

El esquema general del modelo teórico propuesto se muestra en la figura 4. Los componentes de este modelo son:



**Figura 4.** Aplicación de Base de Datos Difusa Sensible al Contexto

- *Usuarios:* Los usuarios a través de los módulos CTP interactúan con el sistema para establecer el contexto explícito: Semántica, Perfil, Rol, Estado (Físico, emocional, social), actividad, situación y/o movilidad. La interacción puede ser efectuada a través de instrucciones del lenguaje estándar del ORDBMS o a través de aplicaciones sensibles al contexto diseñadas para usuarios inexpertos.
- *Módulos CTP:* Consta de los módulos de almacenamiento y recuperación sensibles al contexto (*CAS* y *CAR* respectivamente); éstos son los encargados de comunicarse con el *Gestor de apoyo de Sensibilización al Contexto* para que sean afectadas las entradas (o salidas) de datos de acuerdo al contexto, antes de ser almacenados (o recuperados) los datos en (o de) la base de datos.
- *Gestor de Apoyo de Sensibilización al Contexto (CASM):* es el encargado de gestionar los datos contextuales en el Catalogo Contextual, el histórico de las acciones de los usuarios y las reglas que permitan inferir comportamiento de acuerdo al contexto. Este sistema se convierte en el “Experto” que permite a los *Módulos CTP* hacer su trabajo de transformación.
- *Base de Datos Difusa:* Base de datos que gestiona datos imperfectos siguiendo los postulados de la Teoría Computacional de Percepciones (CTP) de Zadeh (2001), permitiendo la utilización de dominios difusos en los valores de los atributos e incorporando el uso de diferentes métodos (constructores u operadores difusos) para cada tipo de dominio.
- *Contexto explícito:* información de contexto proporcionada por el usuario o a través de alguna aplicación que se conecte al sistema (Semántica, Perfil, Rol, Estado, actividad, situación, movilidad).



- *Contexto implícito*: información de contexto (Ubicación, Tiempo, Clima, Aplicación, Dispositivo, conexión) suministrada por sensores ambientales o agentes de software (servicios web) que infieran el contexto de acuerdo a la actividad del usuario.

Con este modelo pretendemos que las consultas en aplicaciones de Base de Datos Difusa Sensible al Contexto puedan producir resultados diferentes dependiendo del contexto en el cual son ejecutadas. Las características introducidas en nuestro modelo para aplicaciones sensibles al contexto son:

- Representación de datos imperfectos (tanto del mundo real como del contexto) en la base de datos difusa mediante la extensión de datos nativos del ORDBMS para la utilización de dominios difusos.
- El usuario utilizará sentencias del lenguaje de consultas estándar SQL para bases de datos Objeto Relacional, lo cual tiene un doble propósito: utilizar un lenguaje de sentencias muy conocido para los usuarios de base de datos y aprovechar las herramientas nativas que ofrece el ORDBMS para afectar en lo mínimo posible el rendimiento del sistema en general a pesar de ofrecer nuevas facilidades. Para el caso de usuarios que no conozcan el estándar SQL se prevé el desarrollo de aplicaciones amigables para el usuario.
- Se introduce un novedoso elemento en aplicaciones sensibles al contexto: la personalización de la semántica utilizada por los usuarios a través del uso de etiquetas lingüísticas.

A continuación se describirá en más en detalle cada uno de los componentes del modelo.

### 3.5.1. Contexto

Toda la información que gestionemos contiene contexto, así sea establecido explícitamente o sea deducido implícitamente. El conocimiento, gestión y uso del contexto representa uno de los pasos más importantes para abordar el problema de la gran cantidad de información disponible en la actualidad. El contexto se ha utilizado ampliamente como un concepto para dar significado a las palabras en la lingüística y para la comprensión de las circunstancias que rodean un evento particular (Almazan, 2010).

La ciencia de computación ha utilizado el término para circunstancias específicas tales como el cambio de contexto en los sistemas operativos (al cambiar de usuario), presentación de menú contextual en las aplicaciones (al pulsar el botón derecho se ofrecen las opciones disponibles) y las gramáticas libres de contexto de la teoría de la computación. Sólo en las últimas dos décadas se ha intentado elaborar una definición general del contexto, específicamente para la computación sensible al contexto (*context-aware computing*).

El contexto ha sido estudiado en diversos campos de la Ciencia de la Computación entre los cuales destacan: Computación Ubicua, Interacción Persona-Ordenador, Inteligencia Artificial, Inteligencia Ambiental, Recuperación de Información, Sistemas de Recomendación, Computación Nómada y Redes Sensoriales; aunque estas áreas han investigado al contexto con diferentes objetivos, en todas se reconoce que es un vehículo para la realización de sistemas de computación inteligentes. Debido a los efectos que este concepto ejerce en el comportamiento de los seres humanos, se requiere su simulación en los sistemas automatizados para fortalecer las capacidades de comunicar, comportar y procesar información.

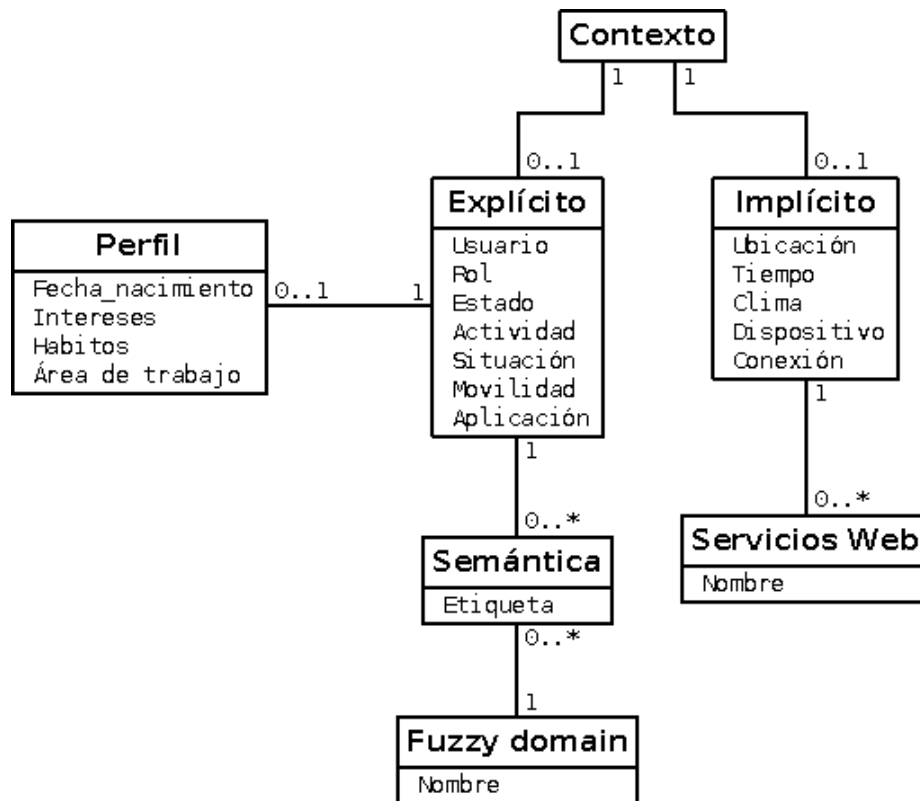
Después de analizar las diferentes definiciones, una de las aproximaciones en el área de HCI es tomar en cuenta al contexto desde el punto de vista de lo que las personas puedan encontrar significativo en la interacción con las aplicaciones, permitiendo hacer más productivas y efectivas las actividades que rodean a los usuarios con respecto a los computadores con el objetivo de incrementar la satisfacción de los mismos, la usabilidad y amigabilidad.

En nuestra propuesta, nos enfocamos en la interacción de usuarios con sistemas de bases de datos difusas, donde se afrontan dos retos que, según nuestro conocimiento, han sido poco abordados: la semántica adaptable de acuerdo al contexto de términos lingüísticos utilizados por cada usuario y la sensibilización al contexto del usuario que interacciona con una base de datos.

Cuando los seres humanos se comunican, hay un contexto implícito que le da un significado a un mismo concepto que puede ser muy distinto. Por ejemplo, la frase *está haciendo calor* puede significar temperaturas muy distintas dependiendo de si la persona está en un país nórdico o un país tropical; ya que el contexto determinado por la ubicación es muy diferente. En el caso de la frase: *quiero escoger una persona de estatura alta*, el significado de la estatura cambia notablemente cuando el objetivo es escoger a un jugador de baloncesto o a un jinete de caballo, aquí el contexto es determinado por los objetivos del usuario. Cuando se dice: *necesito escoger un restaurant que no sea costoso*, la palabra costoso puede significar diferente incluso para una misma persona si el motivo de la cena es para un encuentro familiar, entre compañeros de trabajo o dependiendo de la disponibilidad de dinero de acuerdo al cargo desempeñado; aquí el significado del término puede estar dado por motivos sociales o por disponibilidad de recursos económicos.

Un punto importante es que no se puede desarrollar un modelo general que recoja todas las definiciones de contexto en ambientes de computación, lo interesante es desplegar un modelo flexible que pueda ser adaptado de acuerdo al interés particular, lo cual puede ser proporcionado mediante la utilización del modelado de datos orientado a objeto. A continuación se presenta un modelo del contexto actual del usuario en un esquema de interacción con sistemas de bases de datos con un diagrama UML presentado en la figura 5.

*Definición.* El contexto en un esquema de interacción entre usuario y una Base de Datos Difusa está compuesto a lo sumo de dos objetos: a) Contexto Implícito el cual posee cinco atributos (Ubicación, tiempo, clima, dispositivo y conexión) los cuales pueden ser deducidos de otras aplicaciones o servicios web, b) Contexto explícito el cual posee siete atributos (Usuario, rol, estado, actividad, situación, movilidad y aplicación), éste puede estar asociado a un perfil con cuatro atributos (fecha de nacimiento, intereses, hábitos y áreas de trabajo), de los cuales los tres últimos pueden tener múltiples valores. Además, cada contexto explícito puede estar asociado a definiciones de semántica de términos específicos a través de etiquetas lingüísticas, donde cada una de ellas pertenece a un dominio difuso.



**Figura 5.** Modelo de contexto en un esquema de interacción Usuario con Base de Datos Difusa

Es de hacer notar que este modelo de contexto tiene una clase para la semántica, donde el usuario puede definir diversas etiquetas de acuerdo a los dominios difusos, esto es lo que nos permite innovar con respecto a los modelos de contexto propuestos hasta el momento.

Además este modelo puede ser adaptable de acuerdo a las necesidades del dominio de la aplicación donde se desarrolle el sistema de base de datos difusa sensible al contexto. En el área de modelado contextual los estudios demuestran que aquellos que intentan servir para todas las aplicaciones, al final resultan ser poco efectivos; la aplicabilidad práctica y usabilidad, son parámetros importantes e inversamente proporcionales a la generalidad del modelo, mientras más expresivo y poderoso, será menos práctico y usable. De acuerdo a Bolchini, Curino, Quintarelli, Schreiber y Tanca (2007) debido a la complejidad del problema y la gran variedad de aplicaciones los modelos contextuales, aunque tengan cierto grado de generalidad, deben estar enfocados en un sub-problema específico de contexto.

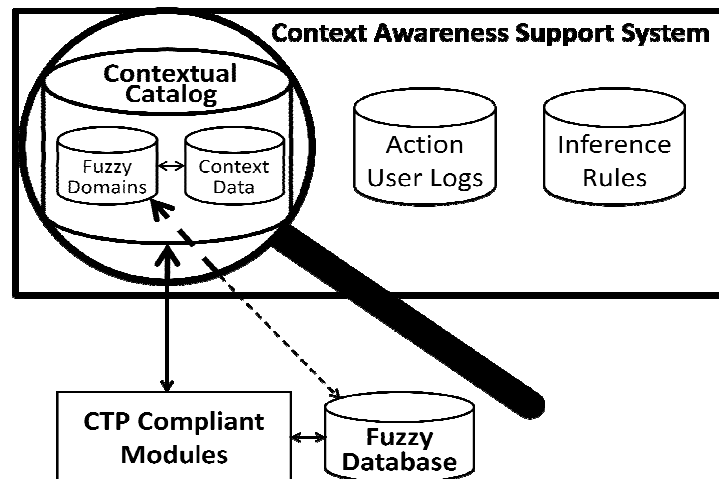
En cuanto a aplicaciones que requieran administrar el contexto dinámico, donde se debe gestionar gran cantidad de datos proporcionados por los sensores, constituyendo cuantiosa información de donde se puede extraer conocimiento (tales como reglas, patrones de comportamiento del usuario), se propone el uso de técnicas como: Bases de Datos temporales, almacenes de datos, flujo de datos y minería de datos, aprovechando experiencias como las de cubos de datos contextuales para datos móviles propuestos por Harvel y col. (2004) o algoritmos de agrupamiento de flujos de datos (*data stream clustering*) como los propuestos en Mahdiraji (2009) y mineras de datos difusos encontrados en Klein, Pery, Komem y Kandel (2000).

La información del contexto será almacenada en un Catálogo Contextual como parte del CASM utilizando la base de datos difusa, lo cual servirá para el desarrollo de aplicaciones y sistemas de bases de datos difusas sensibles al contexto.

### 3.5.2. Gestor de Apoyo a la Sensibilización al Contexto

Tradicionalmente, en los sistemas de bases de datos se ha privilegiado la gestión del contenido, donde el contexto es el establecido por el diseñador del sistema. Se dice en cambio que una aplicación es sensible al contexto, cuando puede decidir qué hacer, basándose no sólo en la entrada explícitamente presentada, sino tomando en cuenta información del contexto, afectando el resultado obtenido de la interacción entre el usuario y el sistema de base de datos.

Es por ello que se necesita un modulo que gestione (modele, almacene, actualice) la información relacionada al contexto. En nuestra propuesta este Gestor lo conforma el Catalogo Contextual, histórico de acciones del usuario y reglas de inferencia; éstas últimas son establecidas a través de funciones o procedimientos del lenguaje estándar del ORDBMS y mediante el diseño de métodos de tipos de datos extendidos (dominios difusos). En la Figura 6, se presenta un esquema del funcionamiento del este Módulo.

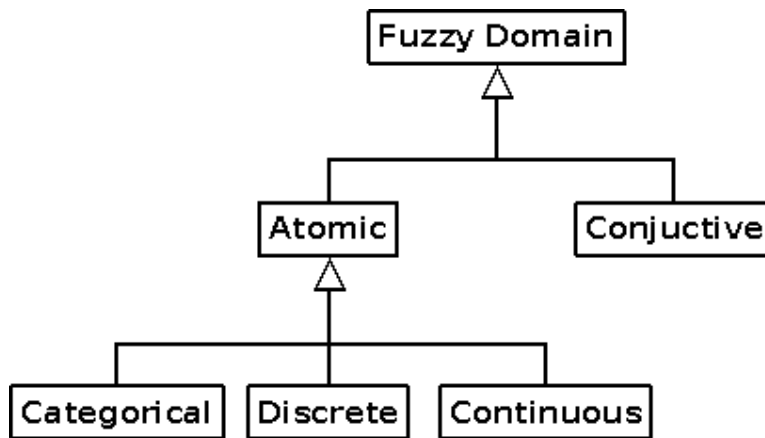


**Figura 6.** Gestor de Apoyo a la Sensibilización al Contexto

Pueden utilizarse módulos de software como los agentes utilizados en la web semántica para extraer o inferir datos contextuales de acuerdo a las acciones que tome el usuario (uso de aplicaciones web, visualización en internet, perfiles de usuario o preferencias establecidas), así como para almacenar en la base de datos un histórico o registro de estas acciones que permita establecer reglas de inferencia de aspectos contextuales.

### 3.5.3. Dominios Difusos

La Base de datos difusa se apoya en la gestión de dominios difusos para modelar, representar y flexibilizar las consultas. En la figura 7 se presenta una jerarquía de clases que incluye los dominios difusos diseñados basado en lo propuesto por Cuevas (2009), utilizando las características orientadas a objeto, extensión de los tipos de datos nativos y creación de métodos utilizando un ORDBMS.



**Figura 7.** Jerarquía de Dominios Difusos

Partiendo de la cardinalidad se pueden distinguir dos tipos de dominios difusos, a saber:

- *Atómico*, son aquellos en los cuales el atributo puede tomar un solo valor que puede ser difuso, a su vez se pueden dividir en:
  - *Categorico*: donde las etiquetas lingüísticas no tienen representación semántica asociada y se crea una matriz de semejanza entre estas etiquetas.
  - *Discreto*: Las etiquetas lingüísticas son expresadas mediante una definición extensiva de un conjunto difuso donde el conjunto soporte tiene una escala discreta.
  - *Continuo*: las etiquetas lingüísticas son definidas por medio de funciones (normalmente lineales) donde el conjunto soporte es continuo (por ej. números reales). En la práctica, generalmente se utilizan funciones (unimodales o monótonas) utilizando varias rectas en su representación gráfica, constituyendo la forma de un trapecio (completo o incompleto) o un triángulo. Estas figuras se representan normalmente por cuatro puntos: A y D (representan el soporte de la función de pertenencia); B y C (representan el núcleo de la función de pertenencia). Debido a la facilidad para su manipulación matemática, el uso de distribuciones trapezoidales como forma de generalizar el resto de funciones es ampliamente aceptado en la comunidad de investigadores de base de datos.
- *Conjuntivo*: permiten que el atributo tome como valor un conjunto difuso que representa un valor múltiple.

Este modelo puede ser usado por diseñadores de aplicaciones de bases de datos, para definir la estructura y el esquema de la base de datos a través de un ORDBMS, o crear interfaces que permitan a un usuario inexperto interactuar con el sistema o para expresar sus percepciones tanto a nivel de definición de etiquetas lingüísticas, como en la obtención de información flexible.

### 3.5.4. Módulos CTP

#### 3.5.4.1. Módulo CTP de Almacenamiento Sensible al Contexto (CAS)

Este módulo permite la inserción de datos de forma amigable a un usuario, mediante la implementación de métodos constructores para los dominios difusos definidos en una base de datos como extensión de tipos de datos nativos de un ORDBMS.

Prevedemos almacenar objetos que tienen una semántica para el usuario que hace la inserción de acuerdo a lo establecido por éste usuario (o en su lugar por un usuario por defecto). De esta forma el CAS sirve como mediador para hacer transformaciones transparentes al usuario conscientes al contexto antes de almacenar la información en la base de datos difusa, de acuerdo a la meta base de conocimiento contenida en el gestor de Apoyo a la Sensibilización al Contexto.

#### **3.5.4.2. Módulo CTP de Recuperación Sensible al Contexto (CAR)**

El Módulo de Recuperación Sensible al Contexto (CAR) permite la recuperación de datos de la base de datos difusa transformados adecuadamente de acuerdo al contexto, utilizando métodos implementados en un ORDBMS y agentes inteligentes programados para tal fin, de acuerdo a la meta base de conocimiento contenida en el Gestor de Apoyo a la Sensibilización al Contexto.

## Referencias

- Aguilera, A. y Rodríguez, R. (2011). Representación y manipulación de datos médicos en marcha patológica. *Revista Multiciencias*, 11(1): 76-84, 2011. Universidad del Zulia, Punto Fijo, Venezuela.
- Ahtinen, A.; Mantyjarvi, J.; Hakkila, J. (2008). "Using heart rate monitors for personal wellness - The user experience perspective," *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE* , vol., no., pp.1591,1597, 20-25.
- Almazan, C. B. (2010). Rover: architectural support for exposing and using context. Disertación doctoral. Universidad de Maryland.
- Atzori, L., Iera, A. y Morabito, G. (2010). The Internet of Things: A survey, *Computer Networks*, Volume 54, Issue 15, Pages 2787-2805, ISSN 1389-1286, 10.1016/j.comnet.2010.05.010.
- Berzal, F., Marín, N., Pons, O., Vila, M.A. (2007). Managing fuzziness on conventional object-oriented platforms. *International Journal Intelligent Systems*;22:781-803.
- Bolchini, C., Curino, C., Quintarelli, E., Schreiber, F. & Tanca, L. (2007). A Data-oriented Survey of Context Models. *SIGMOD Record*, 36 (4), 19 – 26. ACM. New York, USA.
- Bunningen, A., Feng, L., Apers, P. (2006). A context-aware preference model for database querying in an ambient intelligent environment. In: Bressan, S., Kung, J., Wagner, R., editors. *Database and Expert Systems Applications*; vol. 4080 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. ISBN 978-3-540-37871-6; 2006, p. 33-43
- Cadenas, J., Marín, N. y Vila, M. (2011). Fuzzy domains with adaptable semantics in an object-relational DBMS. In: Christiansen, H., Tre, G., Yazici, A., Zadrozny, S., Andreasen, T., Larsen, H., editors. *Flexible Query Answering Systems*; vol. 7022 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg. ISBN 978-3-642-24763-7; p. 497- 508.
- DRAE (2013). Diccionario de la Lengua Española. Real Academia Española. Vigésima segunda edición. Consultado en línea en <http://lema.rae.es/drae/>, el 19-06-2013.
- Jiang, Y., Tang, Y., Wang, J. y Tang, S. (2010) "Representation and reasoning of context-dependant knowledge in distributed fuzzy ontologies", *Expert Systems with Applications*, Volume 37, Issue 8, August 2010, Pages 6052-6060, ISSN 0957-4174, 10.1016/j.eswa.2010.02.122.
- Hadjali, A., Mokhtari, A. & Pivert, O. (2012). "A Fuzzy-Rule-Based Model for Handling Contextual Preference Queries". *International Journal of Computational Intelligence Systems*, 5:4, 775-788
- Harvel, L.D., Liu, L., Abowd, G.D., Lim, Y., Scheibe, C., Chatham, C. (2004). Context Cube: Flexible and Effective Manipulation of Sensed Context Data. ;In *Proceedings of Pervasive*; pp. 51-68
- Hyman, S. E. (2003). Diagnosing disorders. *Scientific American*, 289(3), 96-103.
- Hong, J. Y., Suh, E. H., & Kim, S. J. (2009). Context-aware systems: A literature review and classification. *Expert Systems with Applications*, 36(4), 8509-8522.

- Huynh, K.T., Finance, B. and Bouzeghoub, M. (2012). "Towards an ambient data mediation system", in The 2nd International Workshop on Information Management for Mobile Applications (IMMoA'2012).
- Klein, Y., Pery, R., Komem, J., & Kandel, A. (2000). Fuzzy data mining. In *Intelligent systems and interfaces* (pp. 131-152). Springer US.
- Koutrika, Georgia (2011). "Personalized DBMS: an Elephant in Disguise or a Chameleon?." Data Engineering, 27.
- Koyuncu, M. "Fuzzy Querying in Intelligent Information Systems". T. Andreasen et al. (Eds.): FQAS 2009, LNAI 5822, pp. 536–547, 2009. Springer-Verlag.
- Levandowski, J., Khalefa, M. y Mokbel, M. (2010). An Overview of the CareDB Context and Preference-Aware Database System. *Proceedings of the Very Large Databases (VLDB)*, 3(2):1529–1532.
- Mahdiraji, A. R. (2009). Clustering data stream: A survey of algorithms. *International Journal of Knowledge-based and Intelligent Engineering Systems*, 13(2), 39-44.
- Marín, N., Medina, J., Pons, O., Sánchez, D., Vila, M.A. (2003). Complex object comparison in a fuzzy context. *Information and Software Technology*;45(7):431-444.
- Martinenghi, D. and Torlone, R. (2009). "Querying Context-aware Databases" T. Andreasen et al. (Eds.): FQAS 2009, LNAI 5822, pp. 76–87, 2009. Springer-Verlag Berlin Heidelberg 2009.
- Miele, A., Quintarelli, E., Rabosio, E. and Tanca. L. (2013). A data-mining approach to preference-based data ranking founded on contextual information. *Information Systems* 38, pp. 524–544. Elsevier.
- Oliver, N. and Flores-Mangas, F. (2006). "MPTrain: a mobile, music and physiology-based personal trainer." *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*. ACM.
- Oliver, N. M., & Pentland, A. P. (2000). Towards perceptual intelligence: statistical modeling of human individual and interactive behaviors (Doctoral dissertation, Massachusetts Institute of Technology, Dept. of Architecture).
- Orsi, G. y Tanca, L. (2011). "Context Modelling and Context-Aware Querying (Can Datalog Be of Help?)". O. de Moor et al. (Eds.): Datalog 2010, LNCS 6702, pp. 225–244, 2011. Springer-Verlag.
- Pivert, O., Hadjali, H. y Smits, G. "On Database Queries Involving Inferred Fuzzy Predicates". M. Kryszkiewicz et al. (Eds.): ISMIS 2011, LNAI 6804, pp. 592–601, 2011. Springer-Verlag Berlin Heidelberg
- Pitoura, E., Stefanidis, K. y Zaslavsky, A. B. (2004). "Context in Databases". Technical Report TR 2004-19, Department of Computer Science, University of Ioannina, Greece, 2004
- Roocks, P. Endres, M. Mandl, S. and Kießling, W. (2012). "Composition and Efficient Evaluation of Context-Aware Preference Queries". In *Proceedings of the 17th Int. Conference DASFAA*, Busan, South Korea, pages 81–95. Springer.



- Roussos, I. N., y Sellis, T. (2008) Technical Report. A Model for Context Aware Relational Databases. National Technical University of Athens
- Sadri, F. (2011). Ambient intelligence: A survey. *ACM Comput. Surv.*, 43(4):36:1\_36:66, Oct. 2011.
- Schneider, O. S., MacLean, K. E., Altun, K., Karuei, I., & Wu, M. (2013). Real-time gait classification for persuasive smartphone apps: structuring the literature and pushing the limits. In *Proceedings of the 2013 international conference on Intelligent user interfaces* (pp. 161-172). ACM.
- Stefanidis, K., Koutrika, G. y Pitoura, E. "A survey on representation, composition and application of preferences in database systems". *ACM Transactions on Database Systems (TODS)*, 2011, vol. 36, no 3, p. 19.
- Tanca, L., Bolchini, C., Quintarelli, E., Schreiber, F. A., & Orsi, G. (2011). "Problems and Opportunities in Context-Based Personalization." *Proceedings of the VLDB Endowment* 4, no. 11, 1-4, 2011.
- Wenzel, F. & Soutschek, M. & Kießling, W. (2012). "A Preference SQL Approach to Improve Context-Adaptive Location-Based Services for Outdoor Activities". G. Gartner, F. Ortig (Eds.), *Advances in Location-Based Services*, *Lecture Notes in Geoinformation and Cartography*, pp 191-207, Springer-Verlag Berlin Heidelberg 2012.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*;8(3):338-353.
- Zadeh, L. (2001). A new direction in AI: Toward a computational theory of perceptions. *AI Magazine* 2001;22(1):73-84