

**ENTORNOS VIRTUALES DE INVESTIGACIÓN
PARA LA TOMA DE DECISIONES
EN EL ÁMBITO DEL MEDIO AMBIENTE URBANO**

Antonio José Sáez
César Alcácer
Sonia Rodríguez
Centro Andaluz de Innovación y Tecnologías
de la Información y las Comunicaciones (CITIC)

Artículo Recibido: 05/10/2016

Artículo Aceptado: 14/06/2017

Resumen

Desde los inicios del programa europeo para la investigación e innovación “Horizonte 2020”, la Unión Europea ha promovido el desarrollo y el uso de los denominados entornos virtuales de investigación más allá de los límites académicos habituales. Este tipo de sistemas permiten democratizar las decisiones, investigar y validar resultados y, transferir con mayor celeridad los avances tecnológicos a la ciudadanía. En la presente comunicación se comparte la experiencia reciente del desarrollo de un laboratorio virtual de investigación en materia de gestión de aguas (<https://swirlvre.es>) cuyo diseño conceptual y arquitectura tecnológica pueden encajar en la filosofía latente tras el concepto de *Greencities*.

Palabras clave

Laboratorio virtual, *open data*, toma de decisiones, escenarios, ciencia colaborativa.

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

Abstract

Since the beginning of the European program for research and innovation “Horizon 2020”, the EU has promoted the development and use of so-called virtual research environments beyond the usual academic boundaries. Such systems allow to democratize decisions, investigate and validate results and transfer more quickly technological advances towards citizenship. In this communication we share the recent experience on developing a virtual research laboratory on water management (<https://swirlvre.es>) whose conceptual design and technology architecture can fit into the latent philosophy behind the concept of shared *Greencities*.

1.- Introducción

La gestión e interpretación de los sistemas urbanos, especialmente en relación a la deseada sostenibilidad urbana, requiere de una gran cantidad de datos, de un enorme flujo de información y, de la estrecha colaboración entre el elevado número de disciplinas transversales que intervienen. Pero también necesita de elementos que permitan incorporar la propia incertidumbre asociada a la toma de decisiones y a la implantación de ideas innovadoras que, aunque validadas científicamente, no hayan sido aplicadas aún en el terreno.

Además, la complejidad de los sistemas urbanos se acrecienta cuando se solapan con los sistemas naturales. Esta complejidad radica especialmente en las múltiples interacciones entre la sociedad humana y el medio natural y, más recientemente en los servicios que como urbanitas demandamos de las ciudades y de los servicios ecológicos que necesitamos. La filosofía detrás del concepto de *Greencities* no es más que la de encontrar nuevos enfoques y estrategias alineadas con una demanda social que engloba todo el espectro de actividades socio-económicas dentro de unas ciudades más vivibles.

Para todo ello, se requiere el uso de herramientas específicas. Las decisiones que la sociedad debe tomar y las acciones que son necesarias implantar requieren una información científica de muy alta calidad científica y, del conocimiento y la experiencia de un tipo que actualmente está dispersa, insuficiente o simplemente no disponible. Por lo que es relevante la construcción de un sistema que sea compatible con una mejor comprensión de los sistemas (urbanos/socionaturales) cuyo conocimiento contribuya a una gestión adecuada, así como al desarrollo sostenible e informado, y la toma de decisiones racional. Este tipo de sistemas sirve para apoyar la investigación científica y para la entender la complejidad de este tipo de sistemas.

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

Estos sistemas, que llevan siendo promovidos por la Unión Europea desde hace ya una década, son los laboratorios virtuales o los entornos virtuales de investigación. Las infraestructuras virtuales de investigación proporcionan capacidades avanzadas para los científicos para hacer frente a los grandes interrogantes de la investigación del medio ambiente urbano –en este caso- así como, para hacer frente a los retos científicos sociales y fundamentales urgentes relativos a nuestra vida en las ciudades. Agrega información esencial y la integra en una plataforma analítica que proporciona las herramientas para promover la comprensión y el conocimiento de los procesos que impulsan, mantienen y ponen en peligro los objetivos en las ciudades verdes. Ofrecen, además, nuevas formas de trabajar juntos en el análisis y modelado de datos, simulación y desarrollo de escenarios. Esto contribuye a mejorar la cobertura, la calidad y disponibilidad de la información esencial de *in situ* y observatorios remotos y colecciones.

En la presente comunicación se presenta el diseño y arquitectura de un *Virtual Research Environment* o entorno virtual de investigación (en adelante VRE) que, si bien se ha diseñado para la investigación en materia hidrológica, su esquema responde a las necesidades y requisitos de herramientas que pueden utilizarse en las *Greencities*.

2.- Entorno Virtual de Investigación

Dentro de las directrices en materia de investigación e innovación delineadas en los programas de financiación europeos, en particular en el “Horizonte 2020”, se ha venido definiendo la necesidad de desarrollar *Entornos Virtuales de Investigación* en diferentes ámbitos de trabajo, entre ellos el medio ambiente. Se espera que los VRE den lugar a una colaboración más eficaz entre los investigadores, una mayor eficiencia y la creatividad en la investigación, además de una mayor productividad de los investigadores gracias a un acceso fiable y fácil de identificación, acceso y reutilización de los datos. De esta manera, de acuerdo con dichas directrices, se acelerará la innovación en la investigación mediante el acceso integrado a recursos digitales de investigación potencialmente ilimitados, herramientas y servicios interdisciplinares y comunidades de usuarios. Además, permitirá a los investigadores trabajar y procesar datos estructurados y cualitativos en espacios de trabajo virtuales y ubicuos. Todo ello, contribuirá a una mayor asimilación de la investigación en colaboración y al intercambio de datos entre nuevas disciplinas, comunidades e instituciones de investigación. Por lo tanto, el uso de los entornos virtuales de investigación tiene múltiples ventajas. Por un lado, permite a la comunidad científica alcanzar conocimientos más profundos sobre la biodiversidad, los servicios ecológicos en las ciudades, etc.

Además, los grandes conocimientos aportados por estas herramientas de apoyo a las decisiones facilitan a los cargos directivos la elección de las mejores opciones para gestionar cada uno de los componentes de las *Greencities*. Los avances identificados,

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

implantados y validados, son fácilmente trasladables a otros contextos diferentes pero con desafíos similares.

2.1. Razón de ser

En el plano científico-técnico, la implantación de los laboratorios virtuales supone la evolución lógica (especialmente en el marco europeo) ha dejado de ser entidades monotemáticas adscritas a un departamento universitario cuyo principal objetivo es la producción científica, para ir migrando hacia consorcios internacionales que integran múltiples disciplinas de estudio, formados por equipo transitorios de investigadores (en el marco de un proyecto), constituidos para abordar problemas concretos, y cuyo objetivo es poder materializar la investigación en iniciativas en el terreno no académico.

Por otro lado, nos encontramos con una sociedad que combina una necesidad aviesa de información con la necesidad de gestionar la incertidumbre. Esta incertidumbre deriva del hecho de que el conocimiento se acepta como algo incompleto de, porque (a) se sabe que será superado en algún momento, (b) es socialmente impugnado y (c) no está enteramente controlado por los científicos. En este sentido, el uso los laboratorios virtuales como generadores de conocimiento, pero también como repositorio del mismo, tiene consecuencias muy positivas en la innovación y la aceptación de los avances tecnológicos. Por un lado, aumenta la robustez, entendiendo ésta como la capacidad que tiene el conocimiento de dar respuestas desde todos los ángulos posibles, las disciplinas y las necesidades existentes. Esto viene dado en gran medida porque mejora la solidez y aceptación de los resultados obtenidos, en base a la participación abierta de todos los agentes interesados y el consenso generado a partir de dicha participación conjunta. Por otro lado, el conocimiento científico funciona de manera que satisfaga a los diversos participantes (y los involucrados) en un proyecto de mayor escala. Mientras los avances científico-tecnológicos sólo se vean reflejados en el campo académico o en niveles de madurez tecnológica bajos *Technology Readiness Level* (en adelante TRL) menor o igual a 4, los no versados en materia científica ofrecerán resistencia a la aceptación de “a ciegas” de cualquier propuesta innovadora, algo que los VRE contribuyen a paliar.

Esta robustez dentro de un laboratorio virtual como el Scenario-based Water Innovation and Research Laboratory (en adelante SWIRL) debe verse reflejada por lo menos en los siguientes aspectos:

- Debe ofrecer datos que sean accesibles por múltiples agentes y a diferentes niveles, de carácter multidisciplinar y que sean trazables en el tiempo.
- Debe ofrecerse como plataforma de colaboración científica, para lo cual debe incluir los mecanismos técnicos.
- Debe ofrecer transparencia y replicabilidad de los ensayos, de tal manera

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

que no sólo se puedan comprobar los resultados de las acciones sino el procedimiento por completo.

- Y particularmente, debe desarrollarse de tal manera que favorezca la estandarización y el establecimiento de protocolos.

2.2. Diseño conceptual

Diferentes elementos son necesarios a tener en cuenta en las fases previas de diseño de un VRE. Por un lado, a la hora de implementar estos sistemas existe la necesidad de definir por anticipado las necesidades de gestión, para lo que se recomienda seguir los siguientes pasos:

- Definir una batería de casos de uso, los requisitos del sistema así como la arquitectura de los entornos virtuales de investigación (VREs).
- Integrar de los datos provenientes de las diferentes fuentes relevantes.
- Crear modelos de comportamiento o modelos predictivos.
- Realizar de tests y pruebas de funcionamiento de las herramientas desarrolladas.
- Integrar todas las herramientas en un entorno virtual de investigación.

Por otro lado, una vez definido el alcance de los trabajos, es necesario definir el ciclo de procesamiento de los datos. En este sentido el diseño conceptual del SWIRL (*Laboratorio de Investigación e Innovación basado en escenarios*, por su acrónimo en inglés) se basa en la *Pirámide DIKW* (Datos, Información, Conocimiento y Sabiduría, por su acrónimo en inglés) de obtención del conocimiento. Esta pirámide pone especial énfasis en el cambio de fases, es decir, los pasos que hay que dar para pasar del dato a la información, de la información al conocimiento y del conocimiento a la sabiduría, para su posterior retorno al medio donde se captó el dato.



Figura 1. Pirámide del Conocimiento
Elaboración propia a partir de Rowley (2006)

1. Captación o recepción de los datos. Obtención de los datos a partir de fuentes fiables y en formato estandarizado, fácilmente intercambiable y acompañado de un fichero de metadatos. La fuente es responsable de los aspectos relativos a la medición, mientras que el VRE se ocupa de los protocolos de intercambio de datos.
2. Transformación de datos en información. Los datos han de ser fiables y trazables, pero por sí solos no afloran nada. La información la constituyen los datos procesados, que ya sí tienen un significado por su relevancia, propósito y contexto. Los datos por sí solos son irrelevantes como apoyo a la toma de decisiones. Pero es necesario que sean fiables y trazables (la cadena de custodia). La información, sin embargo, se puede definir como un conjunto de datos que tienen un significado (relevancia, propósito y contexto), que por lo tanto son de utilidad para quién debe tomar decisiones, al disminuir su incertidumbre. Los datos se pueden transformar en información añadiéndoles valor:
 - Contextualizando: se sabe en qué contexto y para qué propósito se generaron.
 - Categorizando: se conocen las unidades de medida que ayudan a interpretarlos.
 - Calculando: los datos pueden haber sido procesados matemática o estadísticamente.
 - Corrigiendo: se han eliminado errores e inconsistencias de los datos.
 - Condensando: los datos se han podido resumir de forma más concisa (agregación).

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

3. Transformación de información en conocimiento. Se produce mediante la comparación de informaciones, la búsqueda de conexiones, el análisis y la reflexión. El conocimiento es una mezcla de experiencia, valores, información y *know-how* que sirve como marco para la incorporación de nuevas experiencias e información, y es útil para la acción (toma de decisiones). Se origina y aplica en la mente de los conocedores. En las organizaciones con frecuencia no sólo se encuentran dentro de documentos o almacenes de datos, sino que principalmente están en rutinas organizativas, procesos, prácticas, y normas. El conocimiento se deriva de la información, así como, la información se deriva de los datos. Para que la información se convierta en conocimiento es necesario realizar acciones tales como:

- Comparación con otros elementos.
- Predicción de consecuencias.
- Búsqueda de conexiones.
- Conversación con otros portadores de conocimiento (avance científico).

Es en esta fase de transformación de la información en conocimiento en los que sistemas de conocimiento avanzado como el *Swirl*, adquieren mayor relevancia. En la gestión de los recursos hídricos prima la modelización (de carácter *ex post* validando los resultados a partir de sucesos previos, respecto a las herramientas de análisis *ex ante*, como las simulaciones o el análisis de escenarios. Los laboratorios virtuales se diseñan con la idea de incluir, tarde o temprano, la posibilidad de realizar modelos *ex ante* y prever la posibilidad de incluir herramientas nuevas en un futuro. En otras palabras, su diseño será lo suficientemente flexible como para permitir que, con el mismo acceso a los datos, se puedan ensayar nuevas técnicas aún en desarrollo, y no sea esclavo de sus propios límites arquitectónicos.

4. Retroalimentación entre conocimiento y experiencia. A través de las simulaciones que se implementen en el laboratorio virtual y que son instrumentos para la evaluación y el aprendizaje.

2.3. La estructura de cubos del conocimiento

La gestión integrada de los recursos hídricos abarca múltiples ámbitos y disciplinas (hidrogeología, sociología, economía de los recursos naturales, ecología...), algo común en comparación con los sistemas urbanos. Tal abanico de posibilidades requiere

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

un planteamiento holístico que permita abarcar todas las disciplinas, pero también una metodología que permita comprobar que todas las necesidades están cubiertas, o identificadas para ser abordadas en el futuro. Al fin y al cabo, la plataforma debe ofrecer una respuesta a distintas necesidades de estudio, adecuarse a los requerimientos de información de los diferentes perfiles de usuario (desde los investigadores a la ciudadanía) y además, debía poder crecer de una manera organizada, con sentido, permitiendo la incorporación de nuevas áreas de investigación y ser escalable.

Para poder casar todos estos requisitos se ideó una estructura que denominamos 'cubos de conocimiento' y que han permitido guiar el desarrollo de las funcionalidades de la herramienta. Se trata de tres ejes, X, Y, Z, cada uno con un contenido específico.

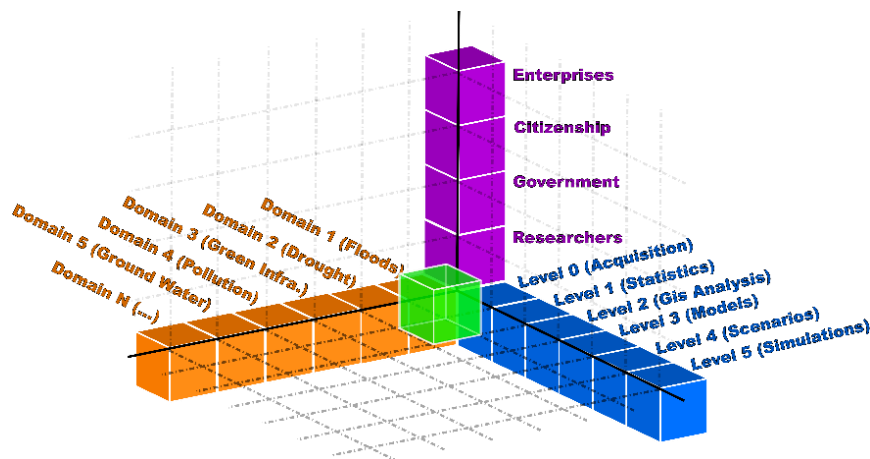


Figura 2. Diseño conceptual de los cubos de conocimiento
Elaboración propia

1. Por un lado, los dominios, que son las áreas en las que tradicionalmente se han encapsulado los datos y que ahora se pretenden poder interrelacionar. Este es un eje abierto a futuras incorporaciones de ámbitos de trabajo.
2. Por otro lado, niveles—un total de cinco—con las posibilidades técnicas que contempla la herramienta y que se construyen de manera agregada, cada uno alimentándose de las capacidades de los anteriores.
3. Y, por último, los usuarios, en primer lugar, los científicos, que es en los que se ha centrado la herramienta en sus inicios; la administración, las empresas y la ciudadanía.

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

4. En cada eje se sitúan una serie de ‘cubos’ a modo de bloques en los que se encapsulan los distintos elementos para poder ordenarlos y priorizarlos. La confluencia de cada uno de estos dominios, niveles y usuarios da lugar a un ‘cubo del conocimiento’ cuyo análisis en profundidad ha evidenciado las necesidades de cálculo y tecnológicas a las que debe dar respuesta la herramienta.

En total se revisaron 15 dominios, que abarcaron desde las inundaciones y las sequías hasta la recuperación de costes o el cambio climático, y se cruzaron con los distintos niveles y usuarios para dimensionar el laboratorio virtual. Esta ‘espina dorsal’ ha permitido clarificar cuáles son las funcionalidades que debe tener el *Swirl* y ha permitido diseñar su estructura desde el punto de vista de la ingeniería de sistemas y software y crear la hoja de ruta para construirlo.

3.- Construcción del sistema

3.1. Estructura de un VRE

Swirl es un entorno virtual de investigación (una plataforma de investigación no física) que se desarrolló con el objetivo de avanzar en el conocimiento de la toma de decisiones en la gestión del agua en diferentes ámbitos: inundaciones, sequías, depuración... Sin embargo, más allá de los aspectos más específicos y característicos de dicha temática (como los informes y cuadros de mandos específicos), la estructura de un VRE es principalmente la misma y por lo tanto replicable.

El objetivo principal de cualquier laboratorio virtual es fomentar el conocimiento y la investigación, no sólo en la temática gestión del agua como es el caso de *Swirl*, sino en especial, en los procesos de toma de decisiones, ya que en la gran mayoría de procesos en los que un laboratorio virtual a nivel de gestión tiene que ver la gestión del riesgo bajo condiciones de incertidumbre. En esta situación, el gestor necesita:

- Apoyarse en la mejor información disponible.
- Hacer uso de las mejores herramientas de modelización que le permitan comprender mejor lo que está sucediendo
- Fomentar el espíritu crítico, plantearse las preguntas correctas y ensayar escenarios.
- Aceptar a todos los actores, entender sus comportamientos y compartir con ellos la información. En definitiva, incluir los modelos sociales en los análisis y la toma de decisiones.
- Prepararse para lo desconocido mediante el entrenamiento virtual.

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

- Comunicar de manera efectiva al resto de agentes y recibir retroalimentación adecuada.

Y todo ello, siguiendo las recomendaciones de la Unión Europea sobre la arquitectura general de los laboratorios virtuales, los cuales se están comenzando a construir en el ámbito europeo con el fin de potenciar un espacio científico único, romper las barreras geográficas con la posibilidad de colaborar *online* y, en consecuencia, acelerar la producción de conocimiento, además de las directivas correspondientes sobre *open data* y acceso a la información pública.

En este sentido, la integración de la información se realiza a través de módulos de ETL (*Extraction, Transformation and Loading*) y también a través de servicios web estandarizados. Estos últimos funcionan como un conector con cada una de las fuentes y trasladarlos hasta el banco de almacenamiento. Así, toda esta información se vuelca sobre un *datawarehouse* un almacén de datos del que beberá el sistema de análisis estadístico, la minería de datos (*datamining*), el análisis geográfico y las herramientas de (*Business Intelligence*) así como los instrumentos relacionados con los modelos hidrológicos. Además, cuenta con bases de datos privadas para que los usuarios puedan incorporar sus propios datos. Para interactuar en el laboratorio los datos integrados son accesibles a través de *interfaces web* para usuarios y de *servicios web* para integración de sistemas SOS (*Sensor Observation Service* o servicio de observación de sensores) y estándares OGC (*Open Geospatial Consortium*). Además, como soporte a los módulos estadísticos y de análisis avanzado le permite desarrollar modelos y ejecutarlos y tiene interfaces específicas para herramientas de simulación de amplia difusión. En cuanto a los resultados, se visualizan a través de informes, cuadros de mando, herramientas de composición de minería de datos y un visor geográfico así como a través de un instrumento específico de *aplicaciones web*.

3.2. Arquitectura de servicios

La gestión del agua aborda muy diversos ámbitos de estudio y, por tanto, la investigación necesita diferentes enfoques. Este laboratorio virtual pretende ofrecer al investigador la posibilidad de abordar cualquier planteamiento. La arquitectura de *software* del *Swirl* se ha construido para ofrecer las máximas prestaciones y para adaptarse a los usuarios en función de su conocimiento y *expertise*. El siguiente es el esquema funcional de la plataforma (se puede clicar sobre él). Se pueden distinguir cuatro niveles o módulos, del 0 al 3, desde la entrada de datos a la creación libre de modelos en laboratorios virtuales (*e-labs*).

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

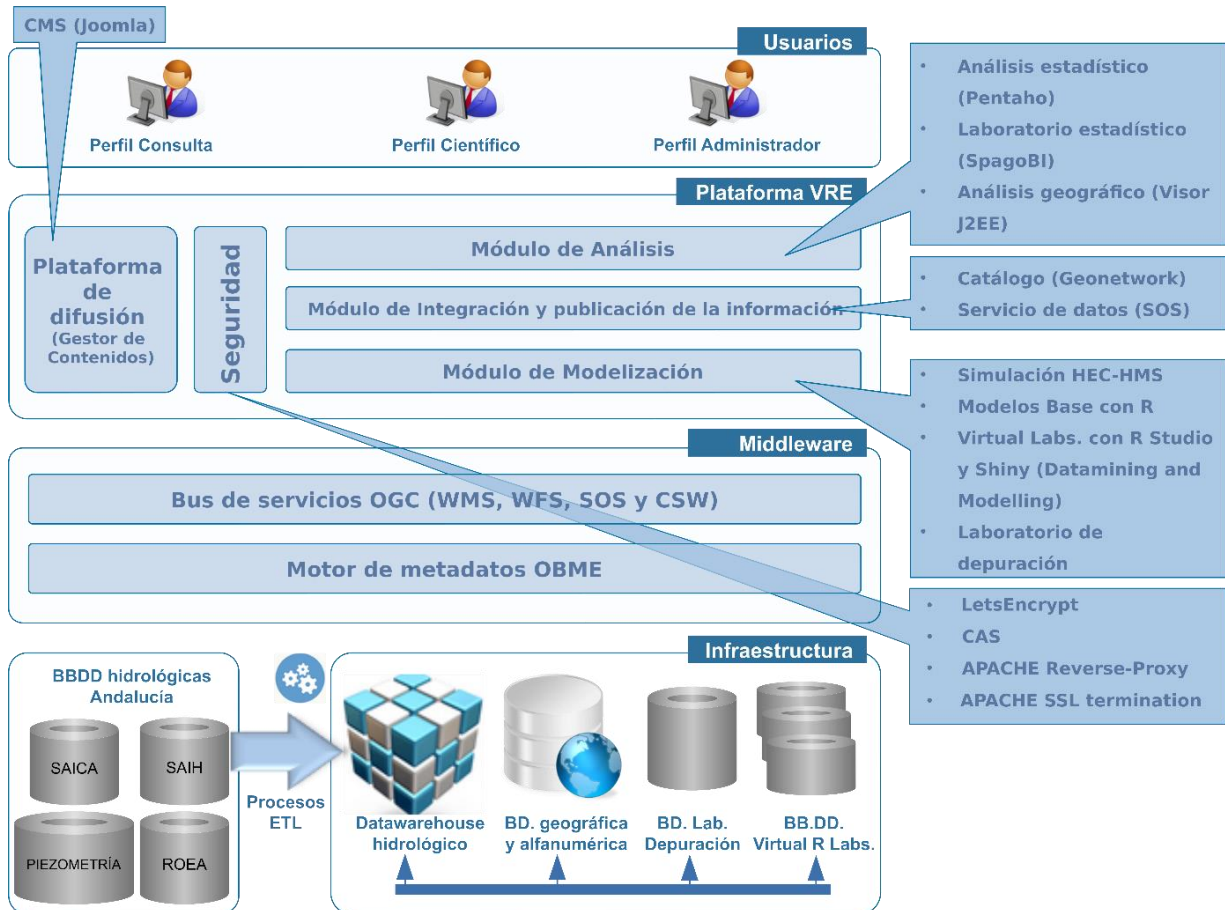


Figura 3. Diseño de la Arquitectura del Swirl
Elaboración propia

A nivel de arquitectura, el sistema se ha diseñado a partir de los siguientes elementos:

Módulo de integración de datos: Este componente tiene como misión recoger la información directamente de las fuentes que ofrecen los datos en continuo (por ejemplo, el Sistema Automático de Información Hidrológica del que disponen todas las Confederaciones Hidrográficas). Esta acción se hace mediante llamadas entre máquinas siguiendo protocolos establecidos, mediante permisos y siempre dentro de una cadena de custodia, en la que no se cuestiona la calidad del dato, sino que se desea tener la trazabilidad del mismo, es decir, saber quién, cómo y por qué ha hecho uso o modificado ese dato.

Data Warehouse (Almacén de datos): Este es el módulo de almacenamiento de los datos. A este módulo tendrán acceso todos aquellos agentes que necesiten acceso

WPSReview International on Sustainable Housing and Urban Renewal (RI-SHUR)

directo al dato. El *data warehouse* es una base de datos que se caracteriza por integrar y depurar información de una o más fuentes distintas, para luego procesarla permitiendo su análisis desde infinidad de perspectivas y con grandes velocidades de respuesta. La ventaja principal de este tipo de bases de datos radica en las estructuras en las que se almacena la información.

Data mining (Minería de datos): Este módulo comprende el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.

Cuadro de Mandos o interfaz del usuario: En un principio será un servicio web, con acceso mediante *login*, en el que se concederán diferentes permisos en función del rol del usuario frente a los datos. Este servicio permitirá la consulta y el manejo de los datos y análisis (algunos prediseñados, otros elaborados *ad hoc*).

4.- Conclusiones

El uso de los laboratorios virtuales puede resultar de gran utilidad en la aplicación de políticas en innovaciones tecnológica en la escala de las *Greencities*. Los VRE representan no sólo un entorno de trabajo para los investigadores, desde donde pueden acceder con facilidad a los datos y compartir en línea con otros investigadores, sino también un punto de encuentro para todos los agentes interesados. Permite el avance científico y su validación, el acceso al dato y sobretodo mejora la toma de decisiones al permitir evaluar opciones a partir de un mismo origen. Sin embargo, el verdadero potencial se encuentra en el uso de este tipo de herramientas para monitorizar la implementación de las acciones en el terreno, mediante el uso de cuadros de mando dinámicos, capaces no sólo de mostrar el estado actual del sistema, sino también de proyectar tendencias a futuro.

5.- Bibliografía

- Allan, R.N. (2009) Virtual Research Environments: From Portals to Science Gateways. Chandos
- Information Professional Series. Woodhead Publishing Limited
- Directiva 2013/37/EU del Parlamento Europeo y Consejo, de 26 de Junio de 2013, como enmienda a la directiva 2003/98/EC para la reutilización de la información del sector público. Diario Oficial de la Unión Europea (27.6.2013)-L175/1.
- Directiva 2003/4/CE del Parlamento Europeo y Consejo, de 28 de enero de 2003, relativa al acceso del público a la información medioambiental y por la que se deroga

WPSReview International on Sustainable Housing and Urban Renewal
(RI-SHUR)

la Directiva 90/313/CEE del Consejo. Diario Oficial de la Unión Europea (14.2.2003)
–L41/26

-Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy.
Journal of Information Science 33 (2): 163-180. doi:10.1177/0165551506070706.