Preguntas y Respuestas Frecuentes

14/05/2019

Dudas sobre el proceso

1. ¿Qué se busca con la Consulta?

Se busca la presentación de propuestas innovadoras destinadas a dar respuesta a una serie de retos mediante el empleo de tecnologías que superen las prestaciones de las existentes actualmente en el mercado.

Estas propuestas servirán para evaluar las capacidades del mercado y definir las especificaciones funcionales que impliquen innovación y sean factibles de alcanzarse a través de una eventual Compra Pública de Innovación u otro instrumento de contratación pública.

El objeto de este tipo de procedimiento no es la recepción de ofertas, sino la recepción de soluciones que ayuden a resolver el reto definido.

2. ¿Cómo se estructura el proceso de Consulta Preliminar al Mercado y cuáles son los próximos pasos al respecto?

Los plazos de la Consulta se rigen por las bases de la Convocatoria, publicada en el perfil del contratante de la Junta de Andalucía y en la página web del proyecto Cloud_IA el pasado 9 de abril de 2019.

- El plazo para la presentación de propuestas dio comienzo el día siguiente al de publicación de las bases en el perfil y finalizará a los 45 días naturales.
- La Consejería de Agricultura, Pesca, Ganadería y Desarrollo Sostenible podrán contactar con participantes concretos para recabar mayor Información sobre su propuesta, aclarar dudas o solicitar demostraciones. Asimismo, se podrán realizer jornadas informativas, reuniones con los participantes, y cualesquiera otras actuaciones de comunicación y diffusion que se estimen oportunas.
- Las reuniones con las entidades podrán celebrarse de forma individual o conjunta con multiples participantes.
- Concluídas las reuniones con las entidades, la Consejería podrá celebrar un evento de conclusiones de la Consulta Preliminar al Mercado.
- No obstante a lo anterior, la Consejería podrá publicar avances y/o actualizaciones sobre el reto de la Consulta durante su transcurso.

3. De acuerdo con plazo regulado en la resolución de la CPM, para asegurar que una propuesta entra en el plazo estipulado, ¿Cuál es la fecha límite?

El plazo finaliza el 24 de mayo a las 23:59 horas. Se recomienda no enviar las propuestas en los últimos minutos.

4. ¿Es posible presentar varias propuestas?

Es posible, tanto presentar una propuesta integral, como una propuesta parcial. Asimismo, se puede plantear una propuesta en conjunto con otra entidad, etc. Lo importante reside en la claridad de la solución propuesta.

5. ¿Se debe rellenar un solo formulario para todos las líneas del reto, o un formulario por cada línea?

Igualmente, podrá enviar un formulario por línea o campo de actividad dentro del reto Cloud_IA, aunque se recomienda, si se tiene capacidad, propuestas integradoras que abarquen el reto completo. Lo importante residirá en el buen entendimiento de la propuesta y grado de detalle.

6. ¿Dónde se remite la Información de la propuesta?

El formulario y la documentación anexa complementaria (si aplica) se enviarán a la siguiente dirección de correo electrónico: cloudia.cpi.cagpds@juntadeandalucia.es Las propuestas se identifican con el acrónimo que quedará claramente expuesto en el asunto del correo electrónico.

Se podrán enviar sucesivas versiones de una propuesta, con el mismo acrónimo, pero cada propuesta enviada sustituirá completamente a la anterior (por ejemplo "acrónimoproyecto_v2 (3,4…n)"). Por ello, la nueva propuesta deberá incluir todo lo que se considere que sigue siendo válido de las anteriores.

En caso de que una propuesta se presente de forma conjunta por un grupo de personas o entidades, deberá emplearse una única dirección de correo electrónico para los efectos de identificación de la propuesta e interlocución con los proponentes.

7. ¿Vincula la entrega de una propuesta para la participación en un futuro proceso de contratación?

Los posibles procedimientos de contratación futuros estarán abiertos a todas las propuestas posibles que cumplan las condiciones establecidas, hayan o no estado ligadas a la consulta preliminar al mercado. La entrega de propuestas en la Consulta Preliminar al Mercado no comportará la generación de incentivos o ventajas para las

empresas participantes a la hora de adjudicar futuros contratos, ni se reconocerá como criterio de adjudicación o como valor ponderable favorable.

8. ¿Cómo se contempla la confidencialidad de la documentación presentada?

Los participantes incluirán en la información que faciliten su consentimiento expreso para que el órgano de contratación pueda difundir su participación y las cuestiones y/o soluciones planteadas en el procedimiento de consulta.

No obstante, el órgano de contratación no podrá divulgar la información técnica o comercial que, en su caso, haya sido facilitada por los participantes y estos hubieran designado y razonado como confidencial.

Serán los participantes quienes deban identificar la documentación o la información técnica o comercial que consideran que tiene carácter confidencial. Los participantes podrán designar como confidenciales alguno/s de los documentos aportados como anexos. Esta circunstancia deberá reflejarse claramente (de cualquier forma o al margen) en el propio documento designado como tal.

En ningún caso podrá ser confidencial el formulario de solicitud.

9. ¿La información de las fichas de solicitud en respuesta a la consulta del mercado (no la de los documentos adicionales clasificados como confidenciales), será siempre publicada a todos los licitantes?

Puede ser publicada. Lo normal es que se publique únicamente el breve resumen de la propuesta.

10. ¿La Consulta Preliminar al Mercado tiene criterios de selección para las propuestas enviadas?

La Consulta no aplica criterios de selección. Será en las futuras licitaciones cuando se establezcan estos criterios.

11. ¿Qué soluciones se buscan, más o menos alejadas de mercado?

En principio no se están buscando que estén más alejadas del mercado o más cercanas. Se está buscando que sean innovadoras y que aborden el reto y resuelvan la problemática identificada.

12. ¿Es posible la participación de empresas públicas en esta Consulta?

Por supuesto, cualquier tipo de entidad, nacional o internacional.

13. ¿Una vez enviada la solicitud?, ¿Cuáles son los procedimientos o próximos pasos?

Tras la Consulta Preliminar al Mercado, el órgano de contratación dará publicidad a los resultados de la misma, respetando el principio de confidencialidad.

Si lo estima procedente podrá comenzar la redacción y la tramitación de la/s licitación/es a partir de las ideas de soluciones recogidas como resultado de la consulta.

Dudas sobre el reto

14. La resolución(es) de las rejillas, ¿ya están predefinidas o es algo a decidir (1km2, 5km2, 100km, etc.)? (cuanta más resolución, mayores serán los requisitos de almacenamiento y procesamiento).

No están predefinidas, ni los niveles , ni las resoluciones, pues en el proceso de Consulta todas las opciones están abiertas. No obstante se pueden adelantar una serie de indicaciones orientativas al respecto, que ayuden a dimensionar las propuestas:

- La infraestructura plantea una **aplicación para la construcción de rejillas** que deberá permitir implementar todas las posibilidades de rejillas: orígenes de referencia, formas de las teselas, resoluciones, niveles, anidamiento o no de los diferentes niveles, proyectadas o sobre el elipsoide, selección de sistemas de referencia, etc. Este apartado hace referencia al desarrollo de una herramienta, no implica la generación de rejillas teselares más que a modo de pruebas.

- La parte fundamental de la infraestructura a desarrollar, en relación con la pregunta planteada, conlleva la generación de uno o varios sistemas de rejilla, **sobre los que se deberá realizar la adscripción de las variables de interés medioambiental**. Por las características de los conjuntos de datos a incorporar, y por la finalidad que Cloud_IA persigue, puede estimarse que la rejilla, que responderá sobre un sistema proyectado normalizado según la directiva Inspire, tendrá unos 7 niveles de resolución anidados entre 5 metros y 500 metros, con posibilidad de inclusión de 3 niveles intermedios no anidados (50, 100 y 200 metros).

- Con carácter de prueba se implementará uno o varios sistemas de rejillas globales (elipsoidales) hexagonales, que incorporará unos 5 niveles anidados entre 25 y 1.000 metros, y solo para un reducido número de variables (unas 30).

15. ¿De qué cantidad de datos estamos hablando? (para ayudar a dimensionar la capacidad del cloud/solución).

La infraestructura deberá ser diseñada en función de las necesidades que plantea el reto. A partir de la incorporación de un conjunto de datos básicos para la gestión medioambiental (ver pregunta 16), que en algunos casos, según sus características, tendrán sus correspondientes actualizaciones, se deberán poder atender una variada gama de procesamientos y análisis que, en la misma medida, generará como nuevos conjuntos de datos a incorporar al sistema.

Como punto de partida se estima que la incorporación inicial de unas 300 variables al sistema (las adscritas al sistema de teselas en los niveles y resoluciones estimados) ocuparía unos 222 TB de memoria. Dado que la infraestructura deberá

estar diseñada para un uso intensivo por parte de un amplio número de técnicos de la Consejería y externos, y que durante seis meses se someterá a un periodo de prueba en un entorno real, se entiende que esa necesidad de memoria deberá multiplicarse por 3 o 4.

16. ¿La solución se desplegaría en el cloud de la Junta o la infraestructura cloud ha de ser provista también por el proveedor? (o incluso una solución mixta).

En relación al modelo de nube todas las opciones están abiertas, y será la propia propuesta la que justifique el modelo más adecuado para el reto que se persigue.

17. ¿Hay una estimación de cuántas variables temáticas habría que gestionar?.

Las variables temáticas a incorporar a la infraestructura no están predefinidas, y deberán ser todas aquellas que, por sus características y variabilidad, sean representativas para que el prototipo de infraestructura demuestre su potencialidad para dar respuesta a todas las necesidades que se le plantean.

Como ejercicio orientativo se estima en un total de unas 30 fuentes de información, las que serían las básicas para la actividad medioambiental, de las que se derivarían un total de unas 300 variables a incorporar, inicialmente, al sistema.

18. ¿Cuáles son las fuentes de datos de las que provienen estas capas temáticas? ¿En qué formatos serán provistas?.

Las fuentes y los formatos de las diferentes capas temáticas son variadas, si bien se tratará mayoritariamente, de bases de datos públicas accesibles a través de la red, a la vez que fuentes internas de la CAGPDS. No obstante, aunque en la implementación del prototipo no tengan por qué figurar fuentes de datos sometidas a restricciones de uso, la infraestructura deberá contemplar, en su diseño, la incorporación de conjuntos de datos de carácter restringido, los cuales deberán estar sometidos a los sistemas de control de la confidencialidad adecuados.

19. ¿Los datos provienen de terceros sistemas (ej., Oracle database) con los que hay que interactuar para extraerlos?.

Como se ha señalado en la pregunta anterior, las fuentes y los formatos de las diferentes capas temáticas son variadas, y se extraerán de plataformas de descargas públicas. Habrá que interactuar con esas plataformas para los procesos de descarga, gestionando los formatos originales, que se incorporarán al datawarehouse de la infraestructura, como primer paso para su posterior depuración, normalización y tratamiento. En ese sentido habrá que generar servicios de importación de los datos desde sus respectivas plataformas.

19. Con respecto al constructor Dashboard: se espera que sea una especie de herramienta gráfica (a implementar por el proveedor) que permita (como si fuese un workflow) integrar varios componentes gráficos comunes (visores de mapas, gráficos, etc.) y servicios geográficos para construir aplicaciones específicas?.

El modelo de difusión de los datos de la infraestructura se basa en la generación de clientes que, no serán específicos, sino que podrán ser diseñados para dar soluciones a medida de cada necesidad específica. Se trata de clientes que, además de consumir servicios, pueden recibir los datos en su formato original (datos incorporados a la rejilla), y sobre éstos, podrán ejecutar, en local, funciones gráficas, analíticas y de visualización avanzada. El constructor de dashboards deberá facilitar, de una forma amigable, el desarrollo de herramientas de multivisualización de gráficos avanzados y de multivisualización avanzada de datos geográficos, que podrán estar o no concatenados, y que permita la interacción dinámica del usuario.

20. Con respecto a "gestionar el ciclo de vida de los datos e incorporar nuevos desarrollos": ¿se espera que cambios en las variables temáticas (ya en formato de rejilla) se trasladen a las fuentes de origen? ¿Será necesario mantener histórico de los valores de las celdas para poder hacer comparaciones?.

La infraestructura contempla que, además de las adecuadas actualizaciones, el repositorio se pueda retroalimentar con datos procedentes de procesos de análisis y transformaciones de los datos originales. No se considera que ninguno de estos procesos se trasladen a las fuentes originales, es decir, realizar el viaje inverso de datos en formato teselar a sus formatos originales.

Se deberá mantener el histórico de los valores de las teselas, pues el análisis multitemporal es una de las principales funcionalidades de una herramienta de este tipo, que tiene en la generación de indicadores una de sus principales utilidades.

21. Con respecto a la afirmación del documento de bases de la consulta"el desarrollo de los diferentes procedimientos de adscripción predefinidos, y también herramientas de desarrollo de nuevos métodos de adscripción a las rejillas multiescalares." ¿se podría explicar con mayor nivel de detalle a qué se refiere?.

Un aspecto crítico del desarrollo de Cloud_IA, con una profunda carga de I+D, es la adscripción de las variables de interés a cada tesela, y a su diferente nivel escalar. Así, para cada variable (se ha estimado un potencial inicial de 300), se deberá definir un procedimiento de incorproración del contenido temático a cada nivel del sistema teselar que se defina. Ello conlleva un conocimiento temático profundo de cada

fuente de información, y de sus potenciales procesos de modelización, para que, de las múltipels opciones de adscripción posibles, se escoja la más conveniente para los propósitos de la gestión medioambiental.

Para cada variable, en el caso de variables complejas, o para cada tipología de variable, cuando se trate de variables simples, habrá que desarrollar un método de adscripción propio, y, además, desarrollar un software (microservicio) que realice esta función.

Igualmente se deberá crear una herramienta que permita la adscripción de nuevos conjuntos de variables, no contemplados inicialmente, y que, el usuario, en un futuro, pueda incorporarla al sistema.

22. **¿Es correcto suponer que la administración contratante asume el coste total del desarrollo de la solución que se solicita en la página 4 del formulario de solicitud y ninguno de los costes de operación posteriores ni riesgos de demanda futura?.**

Aunque es el resultado del propio proceso de Consulta Preliminar quien lo determinará, Cloud_IA se considera que se licitará como una Compra Pública Pre-Comercial, en la que se desarrollará un prototipo muy avanzado, aunque no completo, de e-infraestructura, que será implementada, y testado su funcionameinto en un entorno real durante seis meses. En ese sentido, todos los costes de I+D serán financiados por la administración contratante.

En relación a los aspectos de operación futura del prototipo, no constituirían parte del actual procedimiento, y, por tanto, no se estima que se asuman tales costes . No obstante, no se descarta a futuro darle continuidad al proyecto, y proceder a la adquisición del producto, bajo la fórmula de Compra Pública de Tecnología Innovadora, o compra pública convencional. Es en relación con esta posibilidad por lo que se incluyen en el Formulario de solicitud preguntas relativas una futura gestión compartida de los derechos de propiedad industrial e intelectual.

23. En relación con las diferentes fuentes de información a integrar para la recolección de datos originales, de cara a disponer de una referencia de partida para el diseño de la solución, y en concreto de los módulos conectores con estas fuentes y los módulos de almacenamiento de estos datos originales, ¿está disponible alguna recopilación/categorización/estimación en cuanto al número de orígenes de datos, su naturaleza (satélite, LIDAR, IOT, drones, HAPS…), y volumetría?.

En las preguntas 14 y 16 se hace referencia a los requerimientos potenciales de almacenamiento de las fuentes de información y variables a incorporar a los sistemas teselares.

Por lo que respecta a una estimación de los orígenes de los datos y su naturaleza, en la diapositiva nº24 de la *"Presentación Taller Informativo CPM Cloud_IA. Descripción detallada de aspectos Temáticos del proyecto Cloud_IA.pdf"*, disponible para su descarga desde la página web del proyecto y desde la Plataforma de Contratación, se hace una breve relación de aquellas temáticas de los conjuntos de datos actualmente disponibles, y que podrían constituir el núcleo básico del prototipo de e-infraestructura.

24. En relación con los productos generados mediante el uso de la solución, de cara a disponer de una referencia de partida para el diseño de la solución, ¿está disponible alguna recopilación/categorización/estimación en cuanto al número de productos, su naturaleza, y volumetría?.

En la respuesta a las preguntas 14, 16 y 22 se encuentra información referente a esta cuestión.

25. En relación con los diferentes actores (usuarios de la e-infraestructura), ¿está disponible alguna recopilación/categorización/estimación en cuanto a tipos y número de actores?.

No existe ninguna estimación previa del número de actores. Como elementos de referencia, en el reto se define una e-infraestructura compleja desde el punto de vista de su gobierno y administración, pues estará abierta a técnicos internos de la CAGPDS, especialmente de los ámbitos medioambientales, y, a la vez se hace mención a usuarios externos, los cuales tendrán roles distintos. Un aspecto importante es que los usuarios internos y externos, no solo podrán llevar a cabo tareas de consulta y análisis de la información, y de construcción de sus soluciones a medida, sino que podrán llevar a cabo tareas de desarrollo de componentes, que, debidamente administrados, serán susceptibles de ser incorporados al catálogo de componentes de la e-infraestructura.

Otro aspecto a tener en cuenta para dimensionar la propuesta de solución es que la e-infrestructura debe ser escalable, de forma que, en momentos críticos, debe estar preparada para dar respuesta a análisis complejos, que requieran procesos de computación en paralelo, en corto espacios de tiempo. Este tipo de requerimiento, reservado para emergencias, será muy selectivo y puntual.

26. En relación con las diferentes funcionalidades requeridas y los diferentes actores (usuarios de la e-infraestructura) referenciados, ¿está disponible alguna especificación funcional más detallada que clarifique qué tipos de usuarios debe contemplar la solución y qué funcionalidades deben estar disponibles para cada uno de ellos (por ejemplo mediante casos de uso)? Sería especialmente clarificador poder disponer de esta especificación asociada a las diferentes fases del ciclo de vida del dato, incluyendo ejemplos concretos de herramientas informáticas que proporcionasen dicha funcionalidad (o similar) para cada tripleta usuario-funcionalidad-fase del ciclo de vida del dato.

Como se ha señalado en las cuestiones 24 y 25 hay incertidumbre en cuanto al número de usuarios y su categorización, más allá de estar abierta a técnicos y gestores no expertos en Tecnologías de la Información Geográfica . En relación a las funcionalidades, y potenciales casos de uso, además de las que se señalan en los diferentes documentos del reto (servicios de importación de datos, servicios de normalización, generación de sistemas de rejillas, sistemas de adscripción de variables a las rejillas, etc) se deben inferir a partir de las necesidades propias de la gestión ambiental: modelizaciones por integración de conjuntos de datos, selección y estadísticas, geoprocesos, generación de indicadores, generación de geovisores avanzados, generación de dashboards avanzados, etc.

Por lo que respecta a una “especificación asociada a las diferentes fases del ciclo de vida del dato, incluyendo ejemplos concretos de herramientas informáticas que proporcionasen dicha funcionalidad”, la información más detallada la respecto que ha manejado el Comité Técnico de la CPM de Cloud_IA es la tesis doctoral de la Universidad de Cantabria “Data Management in a Cloud Framework: application to the Lifewatch ESFRI”, Aguilar Gómez F. (2017)

https://repositorio.unican.es/xmlui/handle/10902/13073

27. En relación con la gestión del ciclo de vida del dato, ¿está definido algún requisito en relación con las políticas de preservación de los datos (i.e. periodos de retención en función del tipo de activo)?.

La e-infraestructura debe permitir generar indicadores con la componente histórica de cualquier dato. Para ello resulta necesario conservar permanentemente todas las

actualizaciones de cualquier variable a lo largo de toda la persistencia de la e-infraestructura.

28. ¿Está definido algún tipo de requisito de ciberseguridad para la e-infraestructura (i.e. autenticación, autorización, publicación segura de aplicaciones web, APIs RESTful…) y para la información que almacena (i.e. nivel de confidencialidad de información, integridad, privacidad…)?.

La ciberseguridad, como todo el modelo, será flexible, por lo que dependerá de muchos factores, siendo fundamental el modelo de nube (público, privado o mixto).

En cualquier caso hay un requerimiento expreso, y es que se pueda acceder a la e-infraestructura desde clientes externos, lo que obliga a adoptar los niveles de seguridad estándares. Pero el nivel de seguridad y certificación, más allá de adaptarse al Esquema Nacional de Seguridad, estará abierto a qué modelo y nivel de seguridad se proporcionará a los distintos elementos, en función de la arquitectura que se proponga.

29. ¿Hay algún tipo de normativa o regulación a la que la solución propuesta debe adecuarse?.

Cloud_IA debe adaptarse al contexto en que está situada, estando condicionada por muchas políticas, que van a impregnar a todos los elementos de su arquitectura.

Al ser una arquitectura abierta, el número de normativas a las que tendrá que someterse estará igualmente abierto. Un elemento innovador del proyecto es que su arquitectura se adapte a las necesidades que tienen satisfacer, y éstas llevarán en muchos casos implícitas sus propias normativas (ciclo de vida del dato, confidencialidad de algunos conjuntos de datos, especificaciones de datos y servicios de Inspire, etc.).

en cualquier caso sería interesante que el mercado anticipe en las propuestas aquellas limitaciones, barreras, normativas o regulaciones que pudieran surgir y que la Administración debería tener en consideración a la hora de desplegar la solución.

30. En relación con los derechos de propiedad intelectual e industrial (IIPR), en el caso de que la solución que proponga una entidad incluya total o parcialmente una solución desarrollada con carácter previo (o una evolución) cuyos IIPR pertenezcan a dicha entidad, ¿el reparto de los IIPR y los beneficios de explotación correspondientes a la solución completa podría incluir exclusiones relativos con estos componentes previos?.

Se recomienda que los participantes en la consulta propongan modelos de propiedad intelectual que permitan a la administración estudiar la mejor opción para ambos de cara a la futura licitación, en la cual ya deberá regular todo ello.