**Consejo de diseño 162 Aprovecha las herramientas de visualización de datos pero evita la anarquía.**

Las cada vez más populares herramientas de visualización de datos proporcionan un entorno que aman los analistas de los negocios. Proporcionan la habilidad para definir cálculos, y más importante, para explorar y experimentar con los datos. Los productos han innovado más allá de las antiguas tablas, gráficos de barras, gráficos circulares para facilitar a los usuarios la representación visual de sus datos. Los usuarios de negocio incluso pueden crear un panel de control con elementos de exploración agradables y interactivos para compartir con sus colegas.

Me encantan estas nuevas herramientas. Si fuese un analista de negocios, no cesaría en pedir la compra de una de estas licencias a la dirección de mi empresa. Prometería un retorno importante de la inversión ya que las ventajas que podría ofrecer son bastante buenas. Desde la perspectiva de un usuario, una de las características más valiosas de estas herramientas es la habilidad de relacionar datos dispares, incluyendo datos de fuera del data warehouse empresarial o incluso de fuera de la empresa. Desde la nube. Desde Google Analytics. Desde Excel.

Sin embargo, si fuese una persona de IT, sería ambivalente. Por una parte: les daría a los usuarios lo que quieren. Pueden conseguir datos desde múltiples fuentes, combinarlos como deseen, crear sus propios análisis. El autoservicio de BI (ver consejo de diseño 153) pasa de ser un sueño a la realidad. ¡Hurra! Por otra parte: anarquía. Cada analista crea sus propios conjuntos de datos en las herramientas de visualización, nombrando elementos como quieran, enterrando transformaciones y cálculos dentro del código específico para estas herramientas. De nuevo por todas partes encontramos compartimentos estancos de data marts o spreadmarts. Definiciones inconsistentes, nombres inconsistentes y datos inconsistentes trasladados y transformados múltiples veces los mismos pero diferentes.

Apartándonos de la excitante perspectiva analista-friki y dirigiéndonos hacia una menos divertida perspectiva de auditoría y observancia, la dispersión de las restricciones y una profunda integración lógica dentro de las herramientas de la visualización de datos individuales puede transformarse en una pesadilla. ¿Qué pasa si el analista reivindica demostrar una oportunidad de obtener un beneficio de 10 millones de dólares analizando el mercado a través de media docena de fuentes de datos “cosidas” juntas con reglas de integración hechas manualmente en las herramientas de visualización de datos? ¿Estas reglas de integración están documentadas? ¿Dónde buscas en la herramienta para ver qué ha hecho el analista? ¿Otros analistas utilizan exactamente las mismas reglas?

El personal de ventas de estos vendedores de herramientas de visualización de datos venden un sueño para los usuarios de negocio. Les dicen a los usuarios que no tienen que esperar al IT para construir un esquema de estrella-puedes ensamblar la herramienta directamente al sistema fuente o almacén operacional y construir tu estructura de cubo en memoria tú mismo (o tal vez conseguir una asistencia de consulta barata para los bits frikis). Y del corto al medio plazo, es cierto. Pero no hablan de cómo estará este entorno en un año o dos, cuando la descoordinada demanda analítica de la comunidad de usuarios estén tirando de la infraestructura IT. O cuando los usuarios se cansen de mantener el código barroco subyacente en su panel de control de misión crítica e intenten pasarlo a la IT. Sin mencionar las múltiples versiones del verdadero problema que la inevitable consecuencia de tal anarquía. Aquellos de nosotros que llevamos construyendo data warehouses desde hace tiempo hemos visto esta floración de “bases de datos de escritorio” varias veces, en primer lugar con los lenguajes 4GL en 1980, más tarde con la proliferación de los data marts en los 90 y ahora de nuevo con el increíble poder de las herramientas de visualización de datos. Seamos claros: nos encantan las herramientas de visualización de datos. Sólo queremos que los conocimientos empresariales sean entendibles con estas herramientas, consistentes y fiables.

En algunos casos sería posible implementar toda tu solución BI en una herramienta de visualización de datos, saltando el costoso paso de planificar, diseñar y construir primero una data warehouse. No es ciertamente una buena idea. Un sistema de data warehouse/ inteligencia de negocios del estilo Kimball es más que un simple esquema de estrella. El componente más importante del método Kimball es focalizar en la conformidad: vamos todos juntos y pongámonos de acuerdo con los nombres, definiciones y en las técnicas de cambio de atributos para cada elemento. Definámolos una vez, usemos una herramienta real ETL para planificar e implementar ese movimiento de datos y almacenemos esos datos transformados en un formato duradero (el data warehouse).

Si la anarquía no es la respuesta, ¿qué recomienda Kimball group? Primero, encuentra un lugar en tu arquitectura para una herramienta de visualización de datos. Tus usuarios de negocio lo quieren, y si la organización IT no planea una arquitectura efectiva, encontrarás una arquitectura inefectiva evolucionando bajo tus pies.

Una arquitectura efectiva depende de un data warehouse Kimball, de la gestión de datos implicada y requerida por las dimensiones conformadas. Los usuarios de negocio deben ponerse de acuerdo con el concepto de que cada cosa tenga un nombre, y hacer de ello una parte integradora del modo en que se comunican y hacen negocios. Esto es un problema empresarial, no un problema IT, aunque el IT debe procurar permitir una gestión de datos efectiva. Con cualquier tipo de autoservicio de BI, no es posible prohibir a los usuarios cambiar el nombre o los cálculos de un objeto una vez está en sus análisis, así que necesitamos ayudarles a entender porque no deben hacerlo.

Con un buen diseño e implementación de un data warehouse Kimball en el negocio, los usuarios pueden enganchar las herramientas de visualización directarmente con el data warehouse relacional o con los cubos gestionados de forma centralizada. Las herramientas se usarán sobre todo como implica la etiqueta: para análisis y visualización. Algunos usuarios necesitarán traer datos externos- siempre hay algo que no ha sido adquirido todavía por el data warehouse central –pero debe estar en la periferia. El IT necesita mantener una buena comunicación con los usuarios para ver los datos externos que están utilizando, y evaluar si deben adquirirse de forma centralizada. De manera similar, sería apropiado promover algunos cálculos definidos por los usuarios en la gestión centralizada del cubo, o la capa semántica.

Si tu data warehouse está planificada pero no está en marcha todavía, necesitas dar apoyo al uso interno en contra del almacenamiento de datos operacional o incluso del sistema fuente. Funcionará mejor en el largo recorrido si ayudas proactivamente a la comunidad de usuarios a utilizar las herramientas de visualización de manera efectiva. Creando una asociación, ganarás perspectiva en cómo están utilizando las herramientas. Puedes incluso ganar algo de influencia sobre lo que hacen y cómo lo hacen. Debes ser muy claro al respecto de que cualquier cosa que se haga directamente en el sistema fuente o ODS puede provocar la re-factorización del data warehouse una vez está disponible.

Finalmente, hay organizaciones que no tienen un data warehouse, no planifican tener uno, no trabajan en la gestión de datos, o tienen una relación más bien antagónica antes que colaborativa entre la IT y los usuarios. Si esto describe tu organización, el autoservicio de BI será un resultado inevitable.

<http://www.kimballgroup.com/2014/01/design-tip-162-leverage-data-visualization-tools-but-avoid-anarchy/>