

Curso SIG – Diseño de Geodatabases

Contenido

- Necesidad de un diseño.
- Objetivos de un diseño.
- Buenas prácticas
- Etapas del diseño.
- Diseño moderno para aplicaciones
- Proyectos SIG

Necesidad de un diseño

Implementación efectiva de un SIG se consigue a través de un buen diseño de la base de datos.

Para lograr un buen diseño de la base de datos es necesario hacerse las preguntas correctas:

- ¿Cómo se puede implementar la tecnología SIG para lograr las funciones existentes, o cambiar la forma en que se logra un objetivo?
- ¿Qué datos beneficiarán más a la organización?
- ¿Qué datos pueden almacenarse?
- ¿Quién es o debería ser, responsable de mantener la base de datos?

La respuesta a estas preguntas dependen de cómo se entienda la tecnología SIG, y del conocimiento de la organización.

El diseño para la implementación de un SIG es igual que cualquier otro diseño:

- Comienzo en conocer las metas
- Progresivamente aumentando el nivel de detalle a medida que se obtiene más información, y se acerca la implementación.

Es necesario dedicar el tiempo justo al diseño, ya que de no hacerlo el impacto sobre toda la aplicación actual y futuras puede ser muy negativo.

La base de datos y las aplicaciones asociadas no pueden ser tratadas en forma independiente.

Contenido

- Necesidad de un diseño.
- **Objetivos de un diseño.**
- Buenas prácticas
- Etapas del diseño.
- Diseño moderno para aplicaciones
- Proyectos SIG

Objetivos de un diseño

Diseño es el proceso en donde se definen las metas, se identifican, analizan y evalúan las alternativas de diseño, y se determina un plan de implementación.

El diseño provee una foto de dónde estamos, dónde vamos, y cómo ir de un lugar a otro.

Un diseño de base de datos provee una arquitectura para la base de datos, provee una vista que abarca toda la base de datos permitiendo una evaluación global de la misma desde varios aspectos.

Un buen diseño resulta en una base de datos bien construida, eficientemente funcional y operacional que:

- Satisface objetivos y soporta los requerimientos.
- Contiene los datos necesarios pero no en forma redundante.
- Organiza los datos para su acceso múltiple
- Permite varias vistas de los datos
- Distingue aplicaciones que mantienen los datos de aquellas que solo los usan
- Representa, codifica y organiza apropiadamente elementos geográficos

Beneficios del diseño:

- Aumento de la flexibilidad del acceso y análisis de los datos.
- Facilita la implementación de aplicaciones
- Disminución del costo de captura, almacenamiento y uso de los datos
- Facilita y mantiene datos para soportar diferentes usuarios
- Facilita los cambios futuros
- Minimiza la redundancia en los datos.

Contenido

- Necesidad de un diseño.
- Objetivos de un diseño.
- *Buenas prácticas*
- Etapas del diseño.
- Diseño moderno para aplicaciones
- Proyectos SIG

Buenas prácticas

- Involucrar a los usuarios: ellos adquieren la sensación de ser dueños, ustedes ganan conocimiento.
- Hacer diagramas de a poco, por pasos. Es un proceso interactivo e iterativo. Se progresa apropiadamente en función de las necesidades.
- Crear un equipo de gente, expertos que ataquen cada etapa.
- **Creatividad. *Un nuevo proyecto es una posibilidad para aportar una nueva tecnología y procesos.***
- Dividir grandes proyectos en unidades de trabajo. “Project milestones” no menos de cada dos meses.
- Mantener bien presentes en todo momento los objetivos y metas de la organización, basado en los requerimientos reales.

- No agregar detalles prematuramente, hacerlo en la etapa o momento adecuado. Ej: no tratar de definir todas las reglas de validación antes de que la geodatabase se construya.
- Documentar todo, mantener los borradores, hacer diagramas, etc.
- Flexibilidad, el diseño deberá ser cambiado a lo largo de las etapas hasta llegar a un diseño que se ajuste más a la realidad, que sea aceptado por la gente y apropiado para la tecnología.
- Crear un plan para la implementación del modelo, basado en las prioridades de la organización.

Contenido

- Necesidad de un diseño.
- Objetivos de un diseño.
- Buenas prácticas
- ***Etapas del diseño.***
- Diseño moderno para aplicaciones
- Proyectos SIG

Etapas del diseño

- Modelado de la vista del usuario (requerimientos, funciones, etc)
- Definir objetos y sus relaciones (UML)
- Identificar representaciones de entidades
- Ajustarlo al modelo de geodatabase (UML con objetos propios del software de base)
- Organizarlo en “*datasets*” geográficos

Primeros tres se refieren al modelo conceptual, clasificando los elementos basado en los datos requeridos, decidiendo su representación espacial.

Los dos últimos desarrollan el modelo lógico, ajustándolo al modelo de geodatabase implementable por el software de base a utilizar

Modelado de la vista del usuario

Objetivo – asegurar el entendimiento entre los usuarios y los desarrolladores

- Identificar las funciones o requerimientos de los usuarios, metas y objetivos de los mismos desde un punto de vista corporativo.
- Identificar los datos requeridos por dichos requerimientos
- Organizar los datos en unidades lógicas de elementos
- Definir un plan de implementación inicial
- Identificar funciones asociadas a la organización de la corporación.

Identificación de funciones

Trabajar con funciones de negocio y no con unidades (secciones, departamentos, divisiones) de la corporación:

- Son más estables en la corporación, una unidad cumple con ciertas funciones que en el futuro la cumple otra unidad.

Para cada función determinar una descripción general de actividades asociadas a dicha función.

Actividades pueden incluir el manejo de aprobaciones de ciertas actividades en el terreno (padrones), control del uso del suelo, acuerdos de desarrollo para la construcción de infraestructuras.

Se identifican proveedores y consumidores de la información geográfica

Identificar las fuente de datos

Determinar los datos necesarios para las funciones.

Determinar que datos CREA o USA una función.

Diagrama de flujos:

- Datos que SALEN de la función indica que la función es responsable de la definición, almacenamiento, creación y distribución de los mismos.
- Datos que ENTRAN a la función es responsabilidad de otra función. Intercambio es de muchas formas (formularios, datos, lineamientos, pedidos, respuestas)

La pregunta es:

¿Con quién o con qué interactúa esta función y cuál es la naturaleza de su interacción?

Identificar las fuente de datos

Resolver (aclarar) inmediatamente situaciones de:

- Sinónimos
- Funciones que duplican datos

Interactuar con aquellos que ejecutan la función ya que son los que conocen de los datos.

Usuario debe validar documentos y diagramas asociados a funciones y datos.

Registro territorial	
Tipo de datos	Fuente de datos
Parcela	Subdivisiones
Descripción de la parcela	Títulos territoriales
Fotografías de parcelas	Archivo de históricos
Propietario	Gravamen territorial
Servicios	Ingeniería

Organizar los datos en unidades lógicas

Las unidades lógicas o grupos representan sistemas tales como registro territorial, calles, terreno, distribución de agua, etc.

Cada grupo definido es operado por una función que recibe y/o transmite información.

Ej: Grupo que incluye modelo de superficie con datos sobre niveles de lluvia se conecta con otro que maneja o controla una red hidrográfica ya que le provee ciertos datos hidrológicos, que le permite a la función que controla la red determinar el nuevo flujo de agua provocado por las lluvias recientes.

Cada uno de estos grupos debe tener un sistema de coordenadas común, un tipo de topología (planar, de red o ninguno), y generalmente interactúan entre si.

Etapas del diseño

- Modelado de la vista del usuario (requerimientos, funciones, etc)
- *Definir objetos y sus relaciones (UML)*
- Identificar representaciones de entidades
- Ajustarlo al modelo de geodatabase (UML con objetos propios del software de base)
- Organizarlo en “*datasets*” geográficos

Definición de objetos (entidades) y relaciones

entidad = objetos con propiedades comunes.

- Se identificarán y describirán entidades
- Se identificarán y describirán relaciones entre entidades
- Se documentará las entidades y relaciones mediante diagramas basados en UML

La identificación de entidades y relaciones se puede lograr mediante el análisis de frases de tal forma que un SUSTANTIVO comúnmente identifica una ENTIDAD, un VERBO define una RELACION entre entidades.

Una válvula controla el flujo de gas – entidad (válvulas)

Un dispositivo de gas se conecta a uno o más líneas – Relación entre entidades.

Un sistema de gas se compone de dispositivos y tubos – Agregación de entidades para determinar una nueva entidad más compleja.

Un línea de distribución es un tipo de tubo – Subclasificación

Definición de objetos (entidades) y relaciones

Verbos enmascarados como sustantivos dificultan la determinación de relaciones – ej: conexión, descripción, identificación, agregación)

Para documentar adecuadamente las entidades y sus relaciones se utiliza diagramas basados en **UML**

Etapas del diseño

- Modelado de la vista del usuario (requerimientos, funciones, etc)
- **Definir objetos y sus relaciones (UML)**
- **Identificar representaciones de entidades**
- Ajustarlo al modelo de geodatabase (UML con objetos propios del software de base)
- Organizarlo en “*datasets*” geográficos

Identificar representaciones de entidades

Consiste en clasificar las entidades según la forma de representarlas, basadas en la geometría, o solamente atributos.

Considerar si:

- El elemento debe representarse en un mapa
- La forma del EG es importante o no para el análisis
- El elemento es un dato que puede accederse o visualizarse a través de una relación con otro elemento.
- El EG tendrá diferentes representaciones a distintas escalas (ej. Río)
- Texto será desplegado en pantalla o en productos cartográficos.

Identificar representaciones de entidades

Asignación de tipos:

Punto – ilustra ubicación de un elemento de forma muy pequeña como par a verse como un área.

Línea – ilustra ubicación de un elemento de forma muy fina como par a verse como un área.

Area - ilustra ubicación y forma poligonal de un elemento.

Superficie - ilustra ubicación de un elemento como un área, pero también incluye los cambios en la altura (incluye TINs y ciertos rasters).

Raster – representa un área mediante celdas rectangulares (imágenes satelitales, fotografías aéreas, capa de datos continua) y puede ser usada para varios análisis.

Imágenes, fotos, dibujos – cada uno representa un dibujo digital y no puede ser usado para análisis.

Objeto – identifica un elemento para el cual no se necesita ningún punto, línea, o área, para el cual no existe representación geométrica o gráfica.

Etapas del diseño

- Modelado de la vista del usuario (requerimientos, funciones, etc)
- **Definir objetos y sus relaciones (UML)**
- Identificar representaciones de entidades
- *Ajustarlo al modelo de geodatabase (UML con objetos propios del software de base)*
- Organizarlo en “*datasets*” geográficos

Adecuarlo al modelo de GEODATABASE

La idea de esta etapa es determinar como representar los datos en términos de la geodatabase desde el punto de vista de implementación de la misma por el software.

En esta etapa:

- Se determina la representación apropiada de la geodatabase para las entidades
- Asegurar que las clases determinadas en el diagrama UML del modelo de datos son soportadas por el software.

Adecuarlo al modelo de GEODATABASE

¿Cómo se hace esto con ArcINFO?

Si el tipo de dato espacial es un punto:

- Para puntos no conectados, tales como monumentos históricos, es un PUNTO
- Para puntos conectados, tales como una intersección que conecta calles, es un JUNCTION SIMPLE
- Para puntos conectados que tienen topología interna, tales como una planta de tratamientos, es un JUNCTION COMPLEX.

Si el tipo de dato espacial es una línea:

- Para una línea aislada (ej. Alambrado), es una línea
- Para elementos lineales que participan en un sistema tales como una red vial, es una EDGE SIMPLE
- Para elementos lineales con secciones conectadas tales como líneas de servicio de una red de agua, es una EDGE COMPLEJO.

Adecuarlo al modelo de GEODATABASE

Si el tipo de dato espacial es un área:

- Para una área aislada (ej. parque), es una polígono.
- Para áreas de llenado continuo, tales como vegetación, es un polígono con topología plana (es decir no puede cruzar otro polígono sin cortarlo).

Si el tipo de dato espacial es una imagen (fotografía, mapa escaneado, imagen satelital, u otras), entonces es un tipo RASTER

Si el tipo de dato espacial es una superficie:

- Para superficies en donde los detalles del terreno son importantes, se usa TIN
- Para superficies que cubren grandes áreas y que usan los modelos de elevación existentes, usar RASTER.

Si el tipo de dato espacial es un objeto, se usa OBJETO general. No tiene representación geográfica directa, pero se relacionan con elementos geográficos

Etapas del diseño

- Modelado de la vista del usuario (requerimientos, funciones, etc)
- **Definir objetos y sus relaciones (UML)**
- Identificar representaciones de entidades
- **Ajustarlo al modelo de geodatabase (UML con objetos propios del software de base)**
- *Organizarlo en “datasets” geográficos*

Organizar los elementos en DATASETS Geográficos

Identificar y nombrar los DATASETS geográficos que contendrán las entidades.

- Asignación de entidades a Feature classes y subtipos
- Reagrupación EG relacionados en redes geométricas o topologías planas
- Organización feature classes y datasets en geodatabases.

Reagrupar Feature Classes

Definir la estructura de la FC con subtipos y si pertenecen a un feature dataset o van separadas.

Determinar si una entidad es representable por una FC o un subtipo. Menos FC (por elegir subtipos) logran geodatabases performante.

Cuando es necesario crear nuevas feature classes:

- Cuando cada grupo de elementos relacionados requieren de implementación de comportamientos distintos
- Cuando tiene atributos sustancialmente diferentes
- Cuando se requiere distintos privilegios en el acceso para cada grupo de elementos (o features)
- Cuando algunos features deben ser accedidos a través de versiones y otras no (Versioning sobre algunas FC)

Definición de roles topológicos para las FC

- Si el tipo de feature es Edge simple, Junction simple, Edge compleja o Junction compleja, entonces participa de una red geométrica.
- Todas las componentes de una red geométrica deben pertenecer a un mismo feature dataset, lo cual obliga a que tengan una misma referencia espacial.
- Si el tipo de feature es línea o polígono y la entidad va a cubrir un área completa, tales como parcelas, o si se quiere que los EG que se cruzan se corten, entonces deben pertenecer en un mismo feature dataset.
- Para entidades con features simples, también se pueden ubicar en un feature dataset, el cual funciona como contenedor de capas que puede ser similares o tener algo en común.

Agrupar datasets y FC en geodatabases

Consideraciones:

- Si es una gran corporación, con departamentos con responsabilidades sobre muchos features datasets. Las geodatabases pueden definirse acompañando la estructura de la corporación
- Se puede usar cualquier base de datos relacionales comerciales, pero cada una debe servir a una geodatabase distinta
- Si se está trabajando con geodatabases personales, tal vez se debe determinar nuevas geodatabase debido a limitaciones en el tamaño.

<i>entidad</i>	<i>relacionado a</i>	<i>tipo espacial</i>	<i>GDB software</i>
Agua			
Bomba		punto	objeto
Medidor		punto	punto
Caja medidor	Medidor	punto	punto
Válvula		punto	Junction simple
Tubo		línea	Edge compleja
Planta		punto	Junction compleja

Geodatabase

Feature dataset

Object Class

Feature class de punto

Feature class de punto

Red geométrica

Feature class de JS

Feature class de EC

Feature class de JC

Sistema de Agua

Elementos de R. Agua

Bomba

Medidor

Caja de medidor

Red de Agua

Válvula

Tubos

Planta

Geodatabase

Feature dataset

Object Class

Feature class de polígono

Feature class de relación

Feature dataset

Feature class de línea

Feature class de punto

Base Territorial

Padrones

Propietario

Polígono

Pertenencia

Calles

Calle

Luces de tráfico

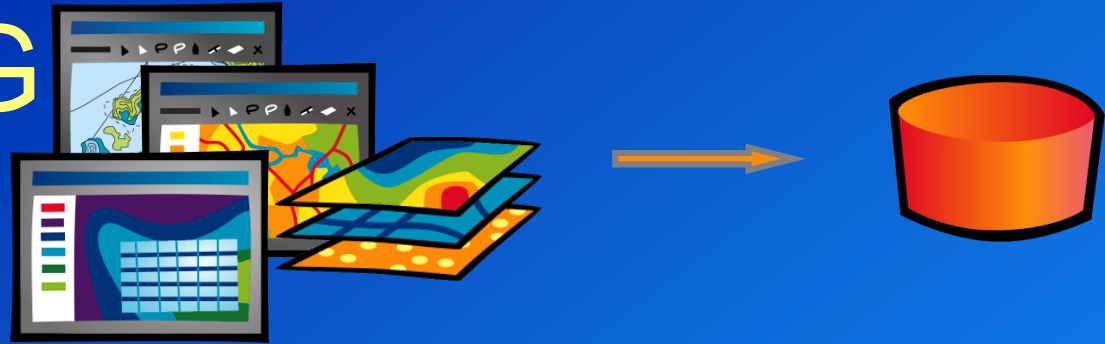
Contenido

- Necesidad de un diseño.
- Objetivos de un diseño.
- Buenas prácticas
- Etapas del diseño.
- *Diseño moderno para aplicaciones*
- Proyectos SIG

Contenido

- Necesidad de un diseño.
- Objetivos de un diseño.
- Buenas prácticas
- Etapas del diseño.
- Diseño moderno para aplicaciones
- **Proyectos SIG**

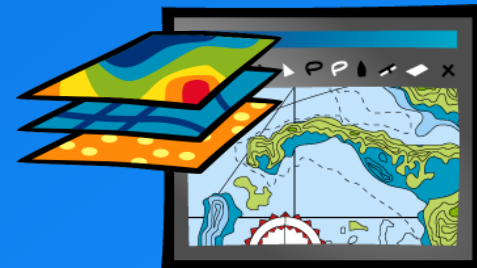
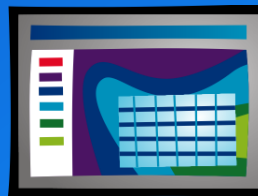
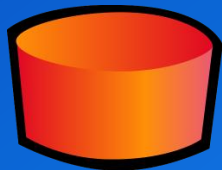
Proyectos SIG



- Puntos de Partida

- Datos existentes, migrar a GDB.
- Nuevas aplicaciones y/o procesos de negocio.
 - Ejemplo: nueva aplicación WEB para el manejo de información catastral.
- Nueva colección de datos y programas de mantenimiento,
 - Ejemplo: Manejo de puntos de direcciones en una ciudad.

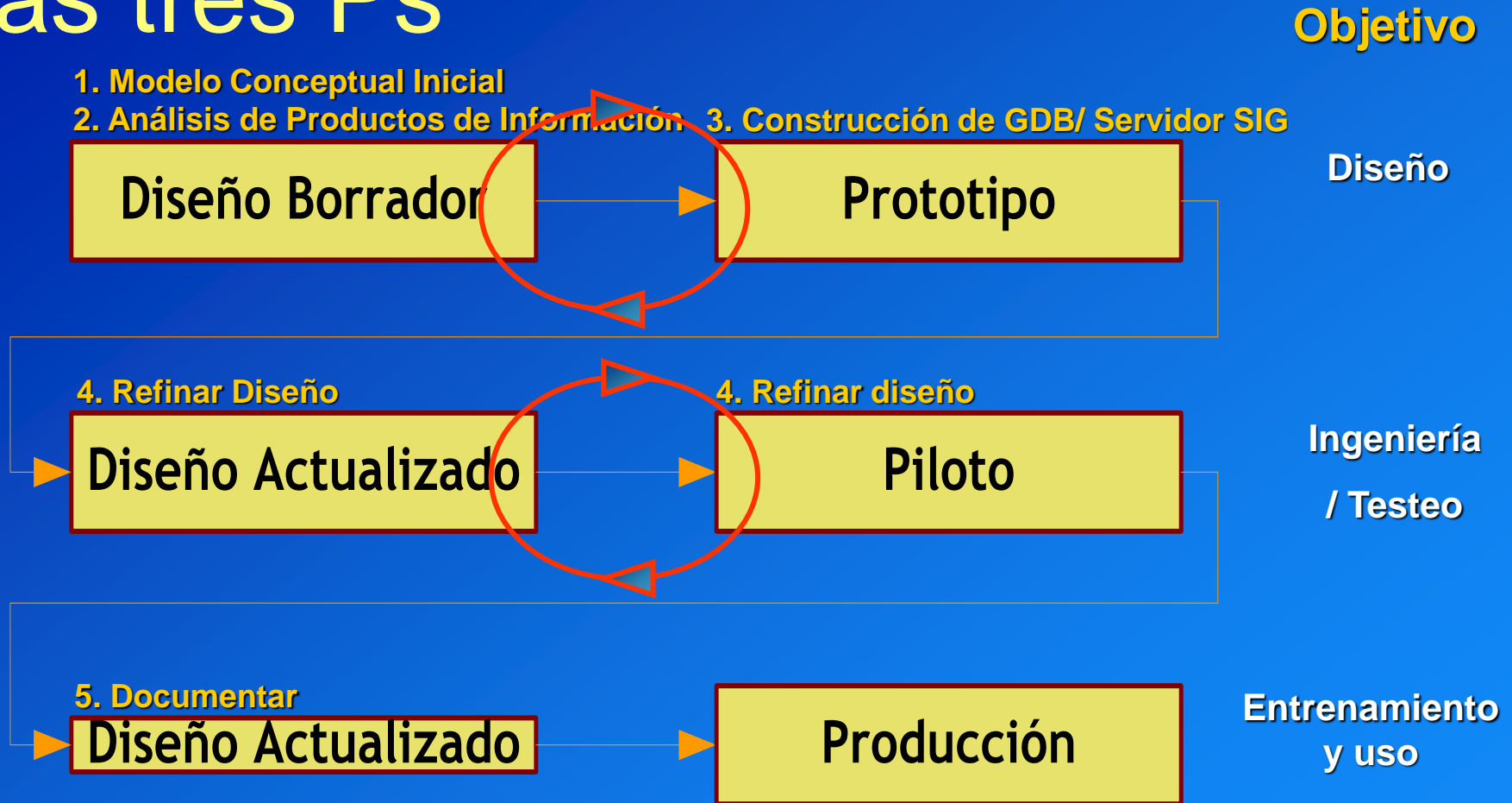
- Consideraciones típicas de un proyecto
 - Focalizar en la aplicación
 - Los requerimientos detectados inciden en el modelo de datos.
 - Focalizar en los datos
 - Experiencia y el conocimiento del negocio incide en modelo de datos.
 - Focalizar en la cartografía
 - Representación de mapas incide en el modelo de datos.



- La mayoría de los proyectos requieren de que por lo menos cumplan con dos características fundamentales.
 - Incluir métodos ágiles/evolutivos
- 3 Estados durante la implementación
 1. Modelo de datos Inicial
 2. Carga de datos, Extracción/Transformación/Carga (ETL por siglas en inglés)
 3. Mapas y Aplicaciones.

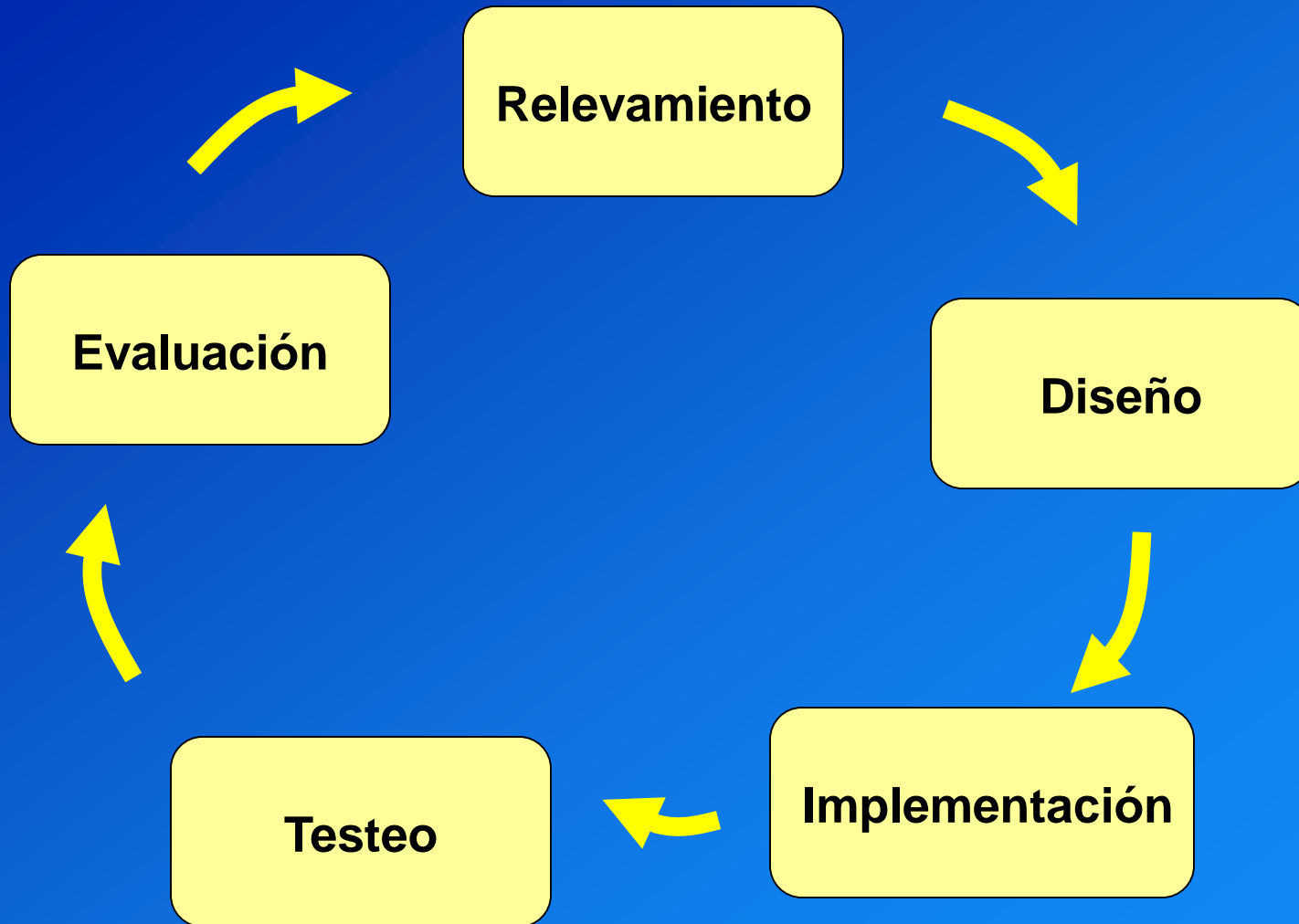
Proceso de implementación

Las tres Ps



Testo de performance y escalabilidad durante la fase piloto es clave para las implementaciones corporativas

Proceso de Diseño



Modelado de Datos

- Tareas Esenciales

- Modelo Conceptual

- Documentar que contendrá la base de datos espacial, como serán mantenidos los datos, como se va a interactuar y como se publicarán.

- Modelo Físico

- Documentar modelo de datos físico en un diagrama UML y codificarlo.
 - Implementar el modelo físico de la GDB.

- Construir una estructura de datos, incluir tablas, relaciones, especificaciones, metadatos, capas, topología, reglas, etc.
- Tareas esenciales
 - Relevamiento de requerimientos
 - Productos (mapas, reportes, etc.)
 - Funcionalidades (análisis, geoprocésamiento)
 - Edición y mantenimiento
 - Metadatos
 - Integración de datos espaciales
 - Performance
 - Análisis y Diseño
 - Crear un modelo conceptual de datosCreate conceptual data model
 - Identificar los datos que se usará para cada proceso.
 - Identificar los datos que se crear a partir de otros datos
 - Crear modelo lógico de datos
 - Identificar datos, metadatos, especificaciones y relaciones.
 - Crear modelo físico de datos, UML
 - Identificar elementos propios de la GDB (feature datasets, tablas, relaciones, dominios, subtipos, redes geométricas, etc.)

- Entregas Claves
 - Documentos de especificación de requerimientos
 - Modelo de datos basados en UML o generados por scripts/código
- Retos y riesgos
 - Aplicaciones tienen una dependencia crítica sobre los modelos entregados.
 - Normalización balanceada.
 - Propagación de cambios.
 - Revisión minuciosa del modelo luego de la puesta en funcionamiento. (usuarios finales).
- ESRI Resources
 - Modelos de datos estándares o disponibles

ESRI Data Models

Address	Health
Agriculture	Historic Perservation
Archiving	Homeland Security
Atmospheric	Hydro
Basemap	Land Parcels
Biodiversity	Local Government
Building Interior Space	Marine
Census	National Cadastre
Energy Utilities	Petroleum
Environmental Regulated Facilities	Pipeline
Fire Service	Raster
Forestry	Telecommunications
Geology	Transportation
GIS for the Nation	Water Utilities
Groundwater	

Normalización vs Desnormalización

- Balance entre el tiempo de adquisición del datos y lo fácil del acceso.
- Desnormalización no es malo pero debe usarse moderadamente.
- Ejemplos
 - Muchos registros relacionados con una sola geometría.
 - Atributos largos que rara vez son editados o accedidos.

