

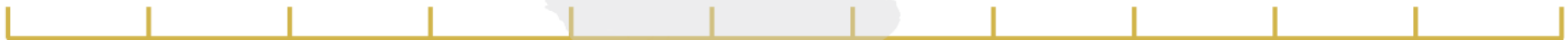
PROMiDAT
IBEROAMERICANO

Programa Iberoamericano de
Formación en Minería de Datos



CRISP-DM

Metodología para el Desarrollo de Proyectos en Minería de Datos



CRISP-DM

CRoss-**I**ndustry **S**tandard **P**rocess
for **D**ata **M**ining

¿Por qué debería ser un proceso estándar?

El proceso de minería de datos debe ser confiable y repetible para personas con escasos conocimientos de minería de datos.



Proceso de Estandarización

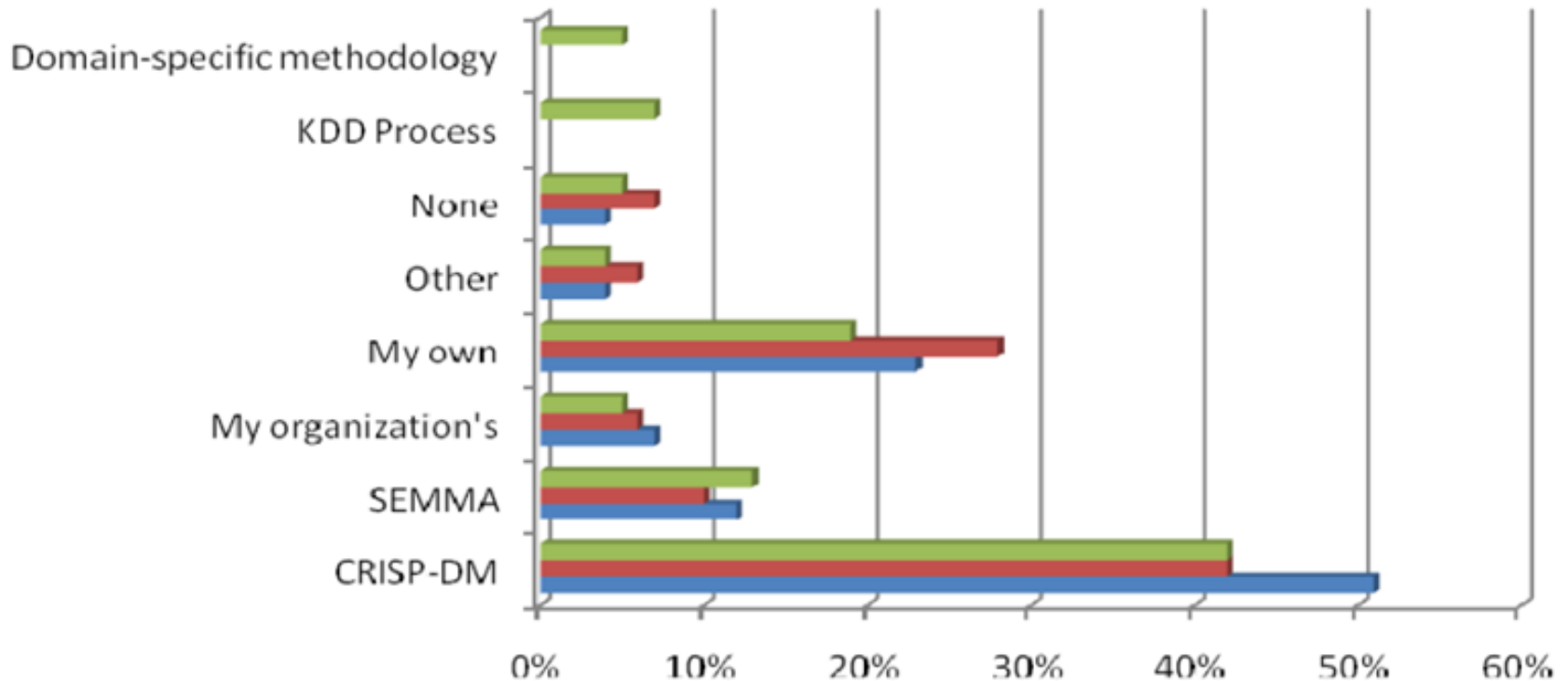
- Iniciativa lanzada a finales de 1996 por tres "veteranos" de la Minería de Datos.
 - Daimler Chrysler (entonces Daimler-Benz), SPSS (entonces COL), NCR
- Desarrollado y refinado a través de la serie de talleres (de 1997-1999)
- Más de 300 organizaciones han contribuido al proceso de modelo de creación.
- Publicado CRISP-DM 1.0 (1999)
- Más de 200 miembros de la CRISP-DM en todo el mundo SIG

CRISP-DM

- No tiene propietario
- Aplicación / Industria neutral
- Se centra en cuestiones de negocios
- Así como en el análisis técnico y de métodos

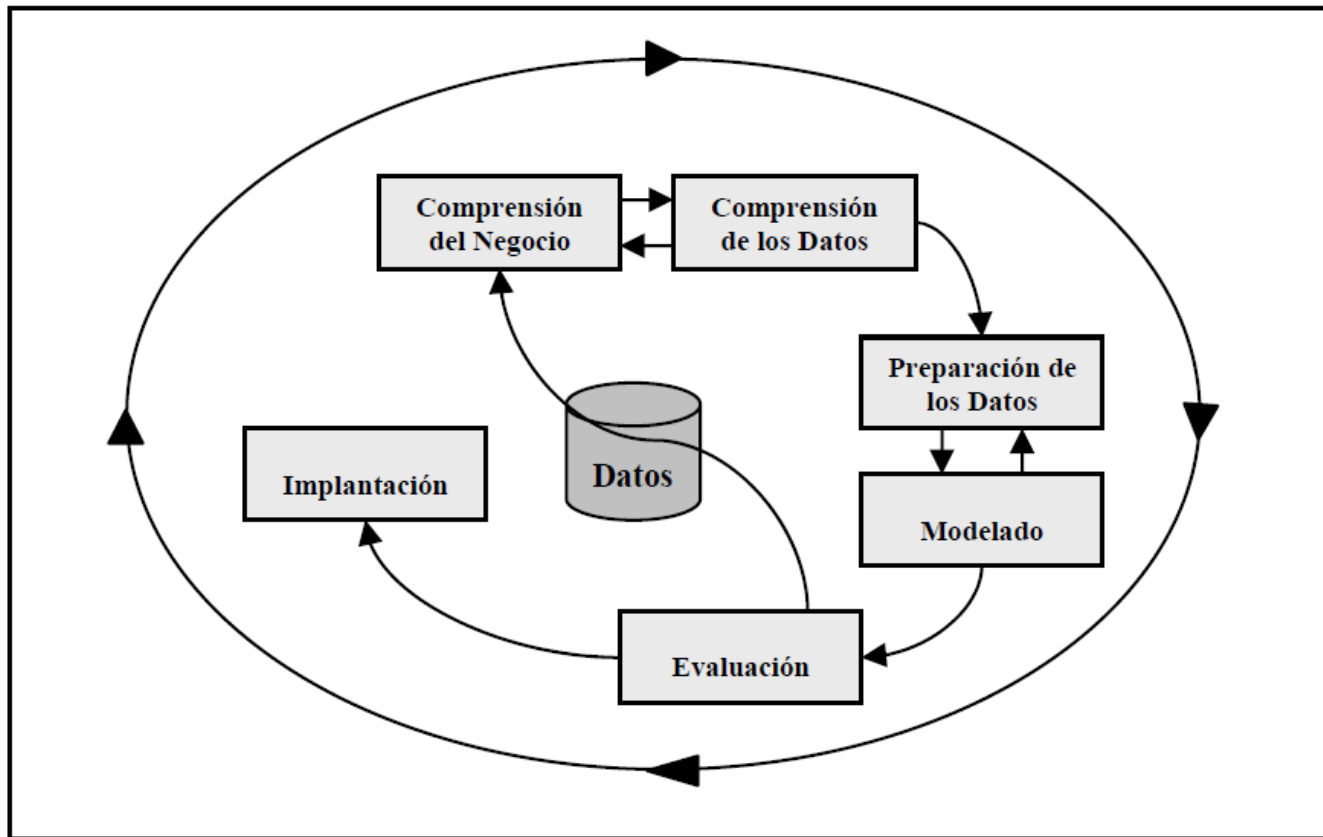


Metodologías utilizadas en Data Mining



Fases de CRISP-DM

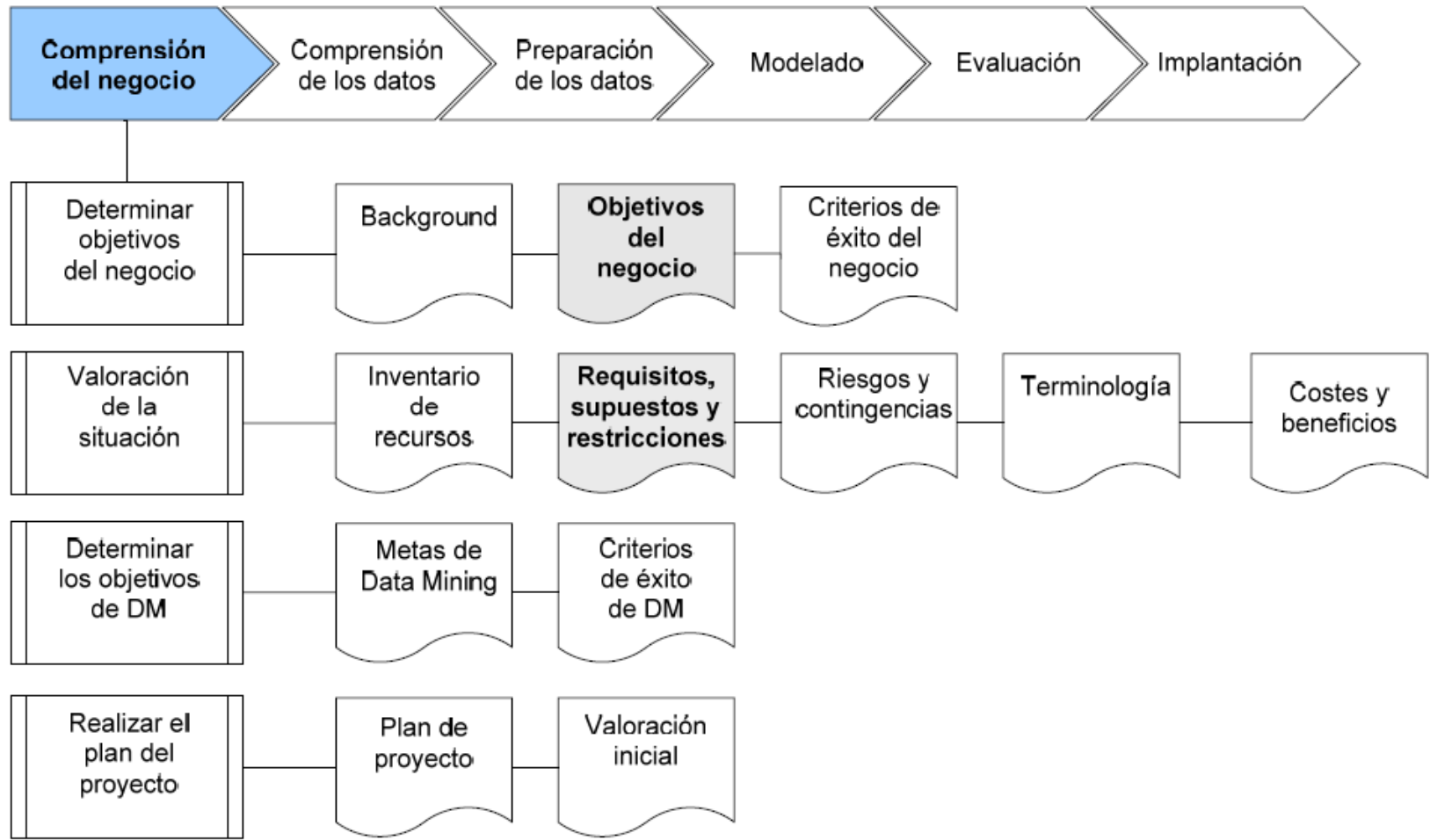
- CRISP-DM, está dividida en una serie de seis fases:



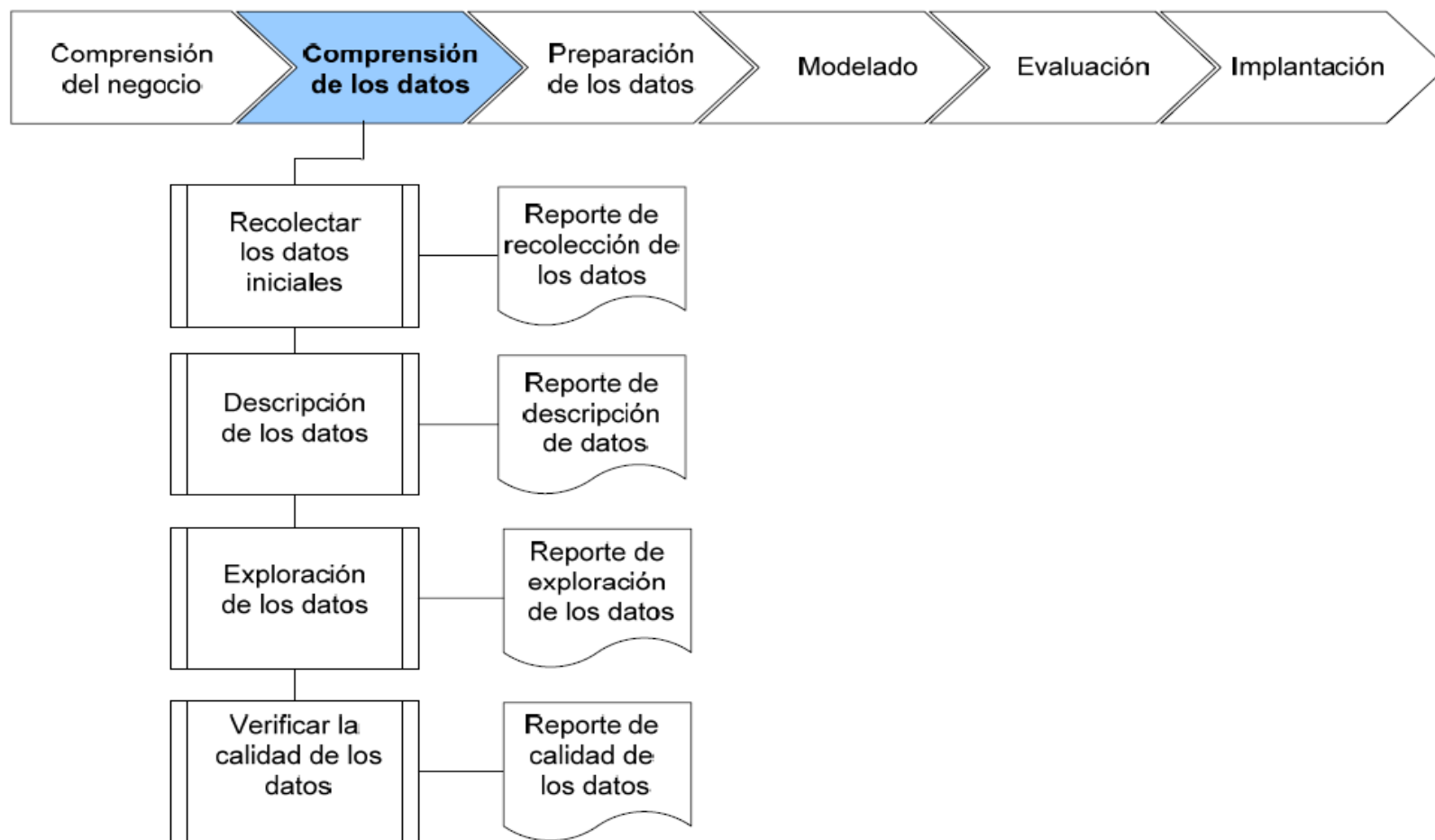
Fases de CRISP-DM

- La sucesión de fases no es necesariamente rígida.
- Cada fase es estructurada en varias tareas generales de segundo nivel.
- Las tareas generales se proyectan a tareas específicas, donde finalmente se describen las acciones que deben ser desarrolladas para situaciones específicas.

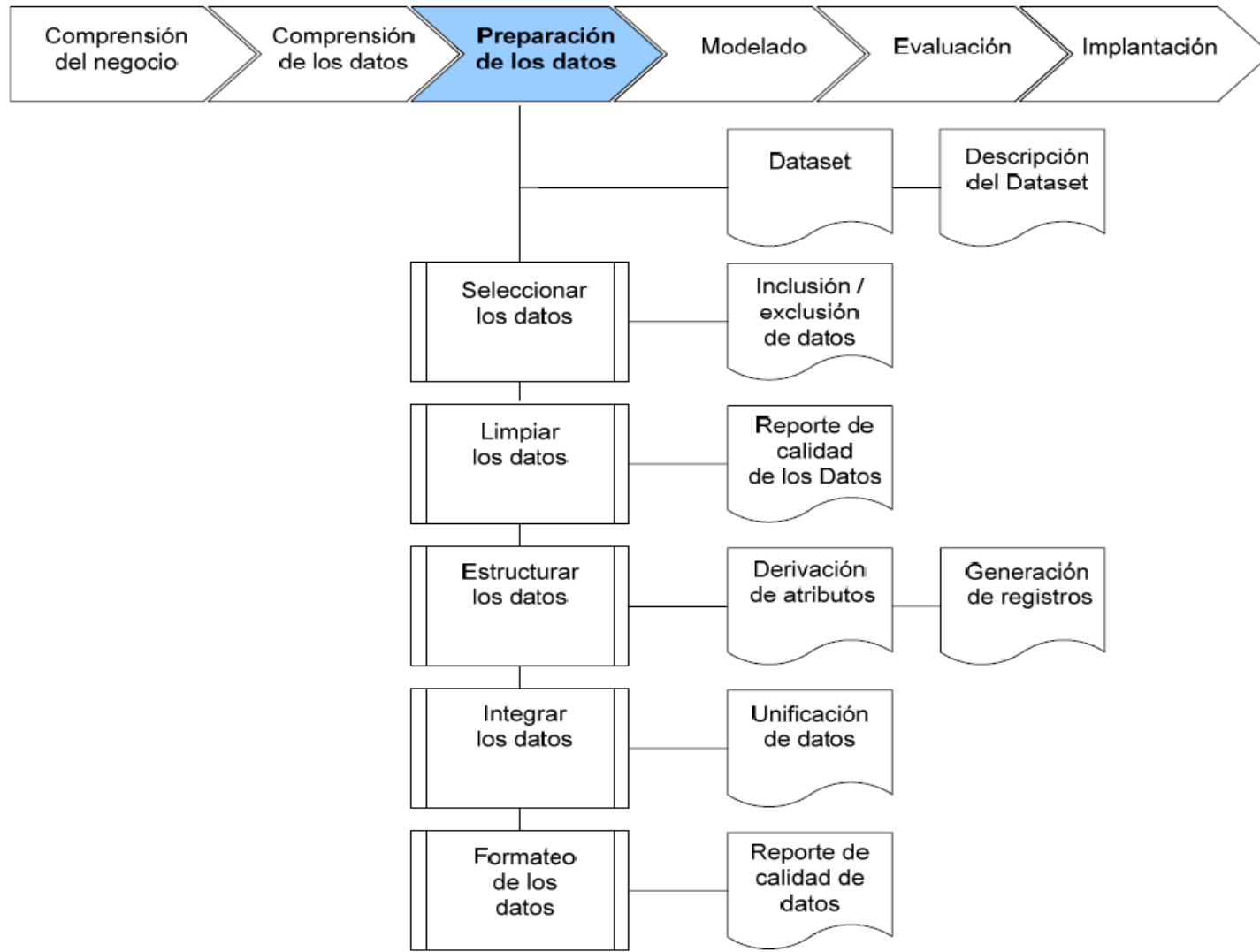
Fase de comprensión del negocio o problema.



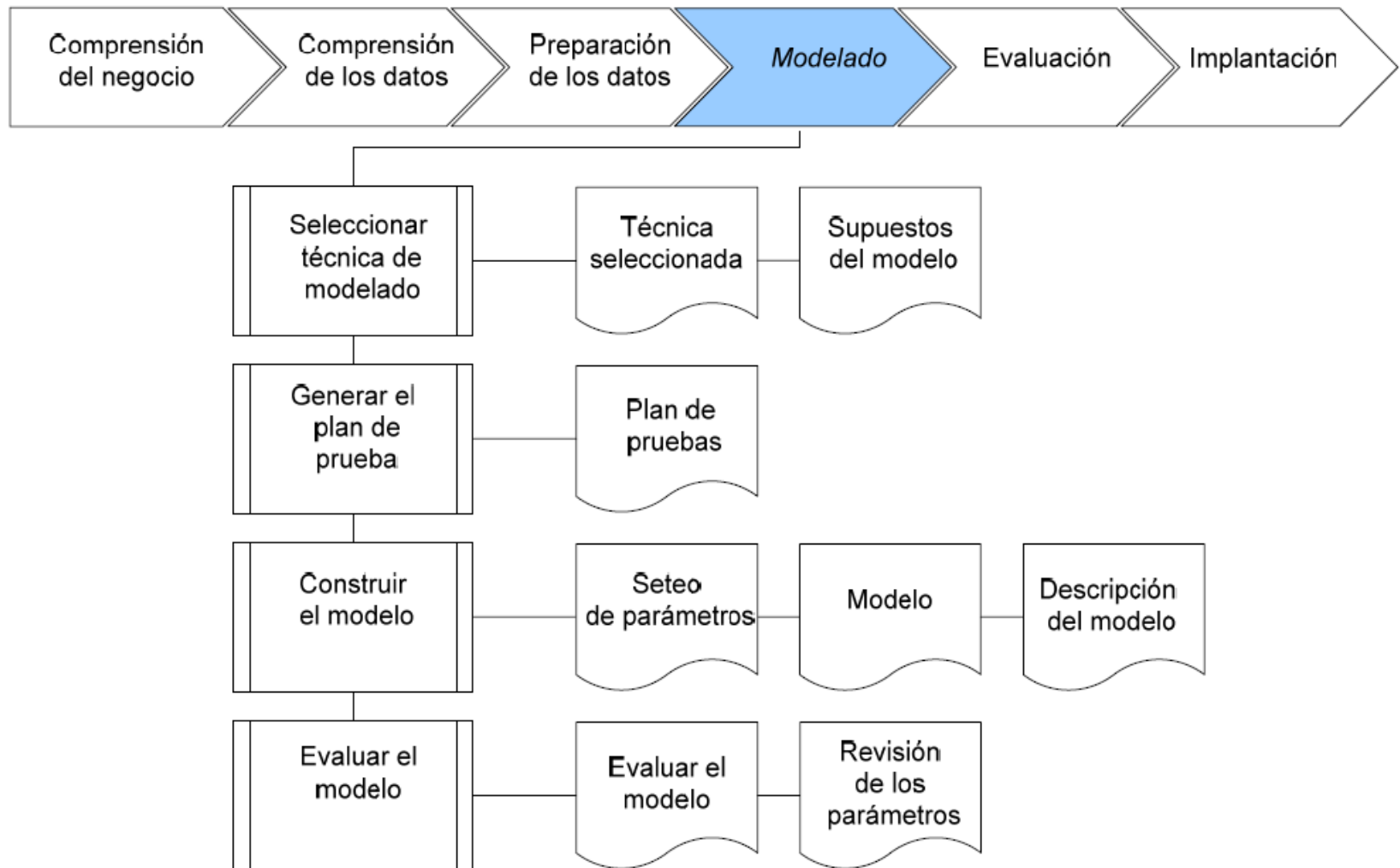
Fase de comprensión de los datos



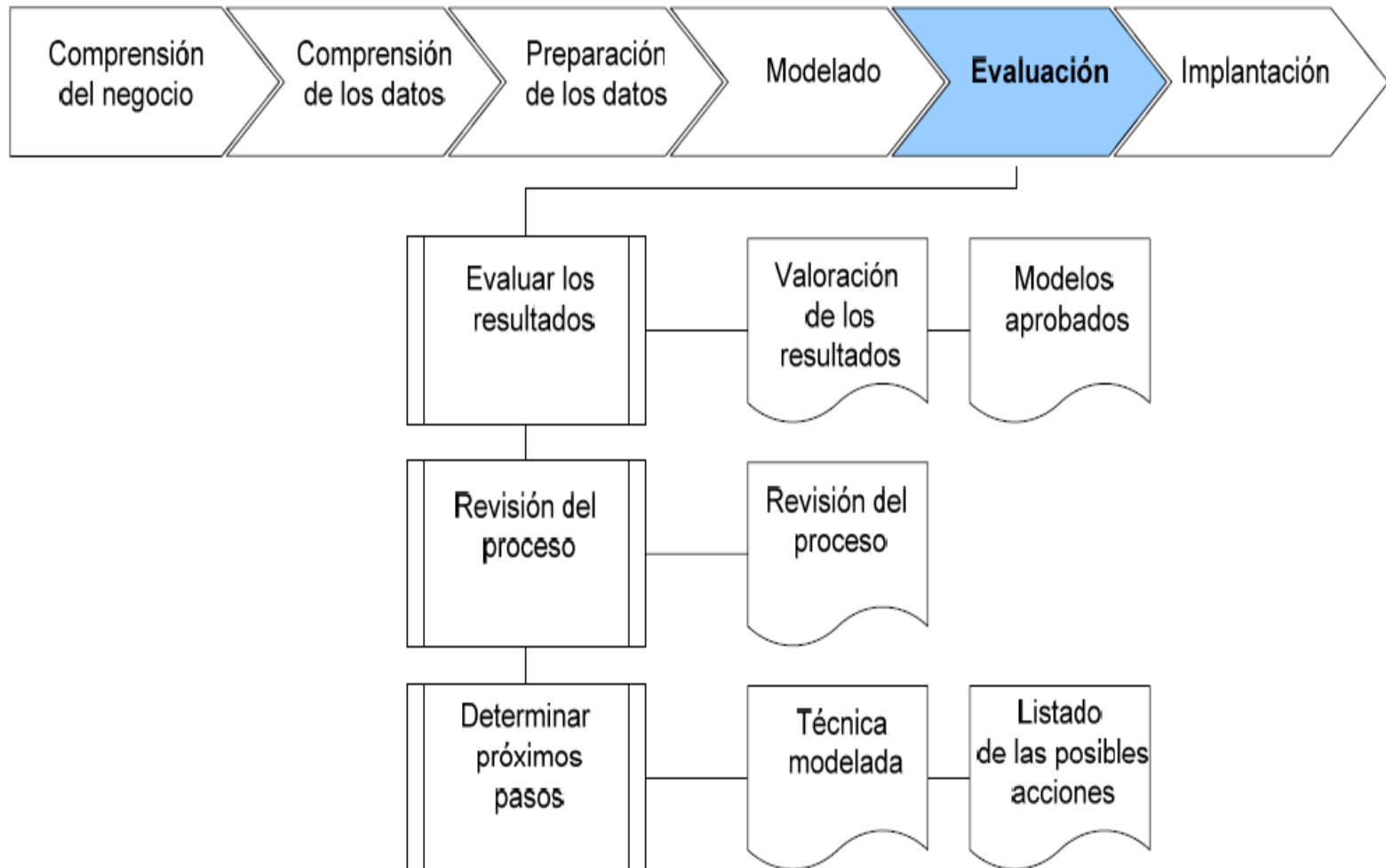
Fase de preparación de los datos



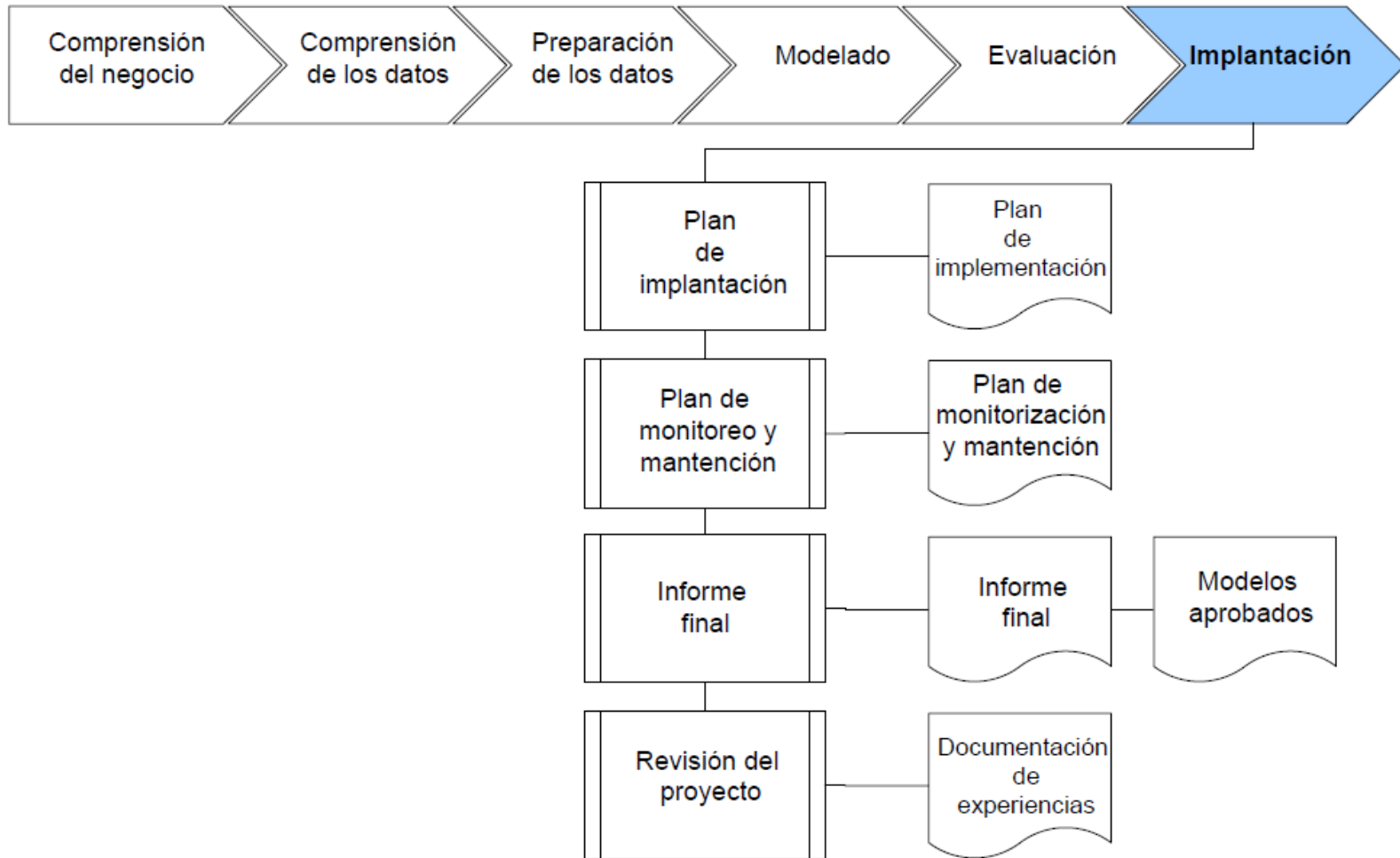
Fase de modelado



Fase de Evaluación



Fase de implementación





PROMiDAT

IBEROAMERICANO

Programa Iberoamericano de
Formación en Minería de Datos

Gracias....