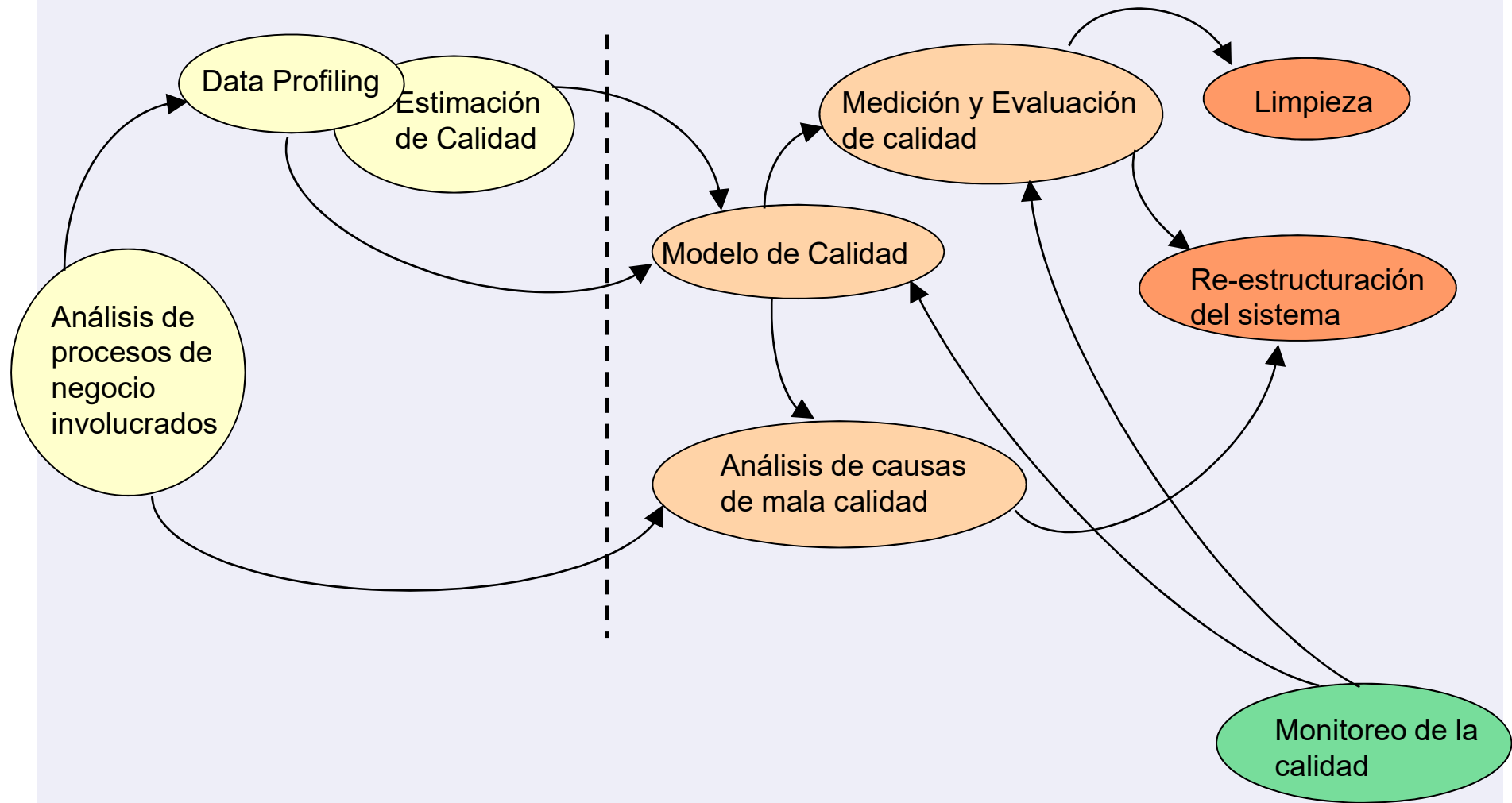

Calidad de Datos e Información

Introducción

Gestión de la calidad en SI



Metodologías de Calidad de Datos

- Una metodología para calidad de datos proporciona un conjunto de pautas y técnicas que, a partir de información de entrada que describe un contexto de aplicación dado, define un proceso para evaluar y mejorar la calidad de los datos [Batini et al., 2009].

Contexto

- Fuente de información poco utilizada en entornos informáticos [Dey, 2001]
- Definición
 - ningún trabajo dice explícitamente qué lo define
 - considerado un concepto abstracto
 - no está claramente definido ni formalizado

Contexto

- Cualquier información que permita caracterizar la situación de una entidad (persona, lugar u objeto), que se considera relevante para la interacción entre un usuario y una aplicación, incluyendo al usuario y las propias aplicaciones [Dey, 2001].
- En Calidad de Datos (CD): Contexto es información de los datos cuya calidad será evaluada (llamados *data at hand*), y otros elementos que están fuertemente relacionados a estos datos.

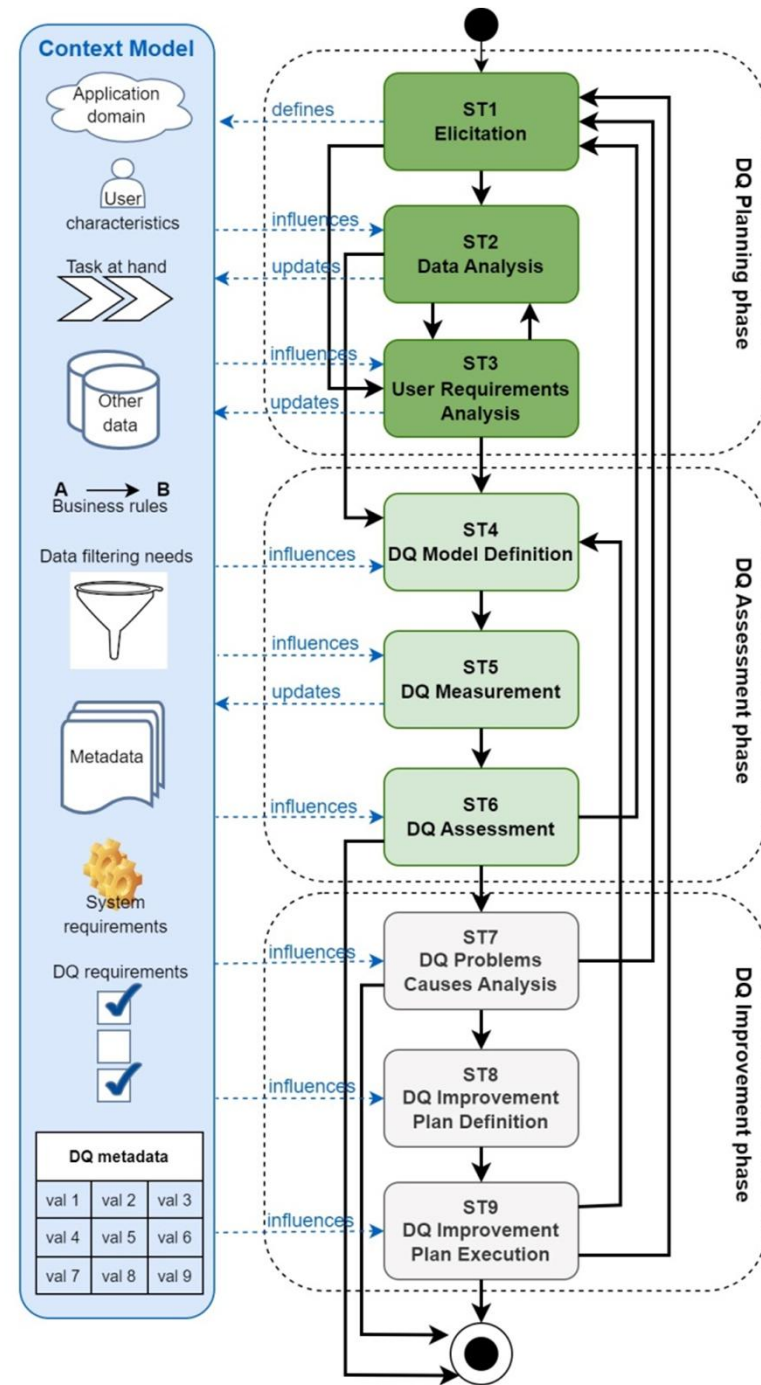
Metodologías de Calidad de Datos

Metodología	Nombre	Año	Autores
TDQM	Total Data Quality Management	1998	R. Y. Wang
TIQM	Total Information Quality Management	1999	L. P. English
COLDQ	Cost-effect of Low Data Quality	2001	D. Loshin
DQA	Data Quality Assessment	2002	L. L. Pipino, Y. W. Lee, R. Y. Wang
AIMQ	A Methodology for Information Quality Assessment	2002	Y. W. Lee, D. M. Strong, B. K. Kahn, R. Y. Wang
CDQ	Comprehensive Methodology for Data Quality Management	2006	C. Batini, F. Cabitza, C. Cappiello, C. Francalanci, P. di Milano
HIQM	Hybrid Information Quality Management	2006	C. Cappiello, P. Ficiaro, B. Pernici
DQPA	A Data Quality Practical Approach	2009	M. Del Pilar Angeles, F. García-Ugalde
HDQM	Data Quality Methodology for heterogeneous Data	2011	C. Batini, D. Barone, F. Cabitza, S. Grega
DQAF	Data Quality Assessment Framework	2013	L. Sebastian-Coleman
TBDQ	Task-Based Data Quality Method	2016	R. Vaziri, M. Mohsenzadeh, J. Habibi
OODA DQ	Observe-Orient-Decide-Act Methodology for Data Quality	2017	A. Sundararaman, S. K. Venkatesan

Ejemplos tomados de [Cichy and Rass, 2019].

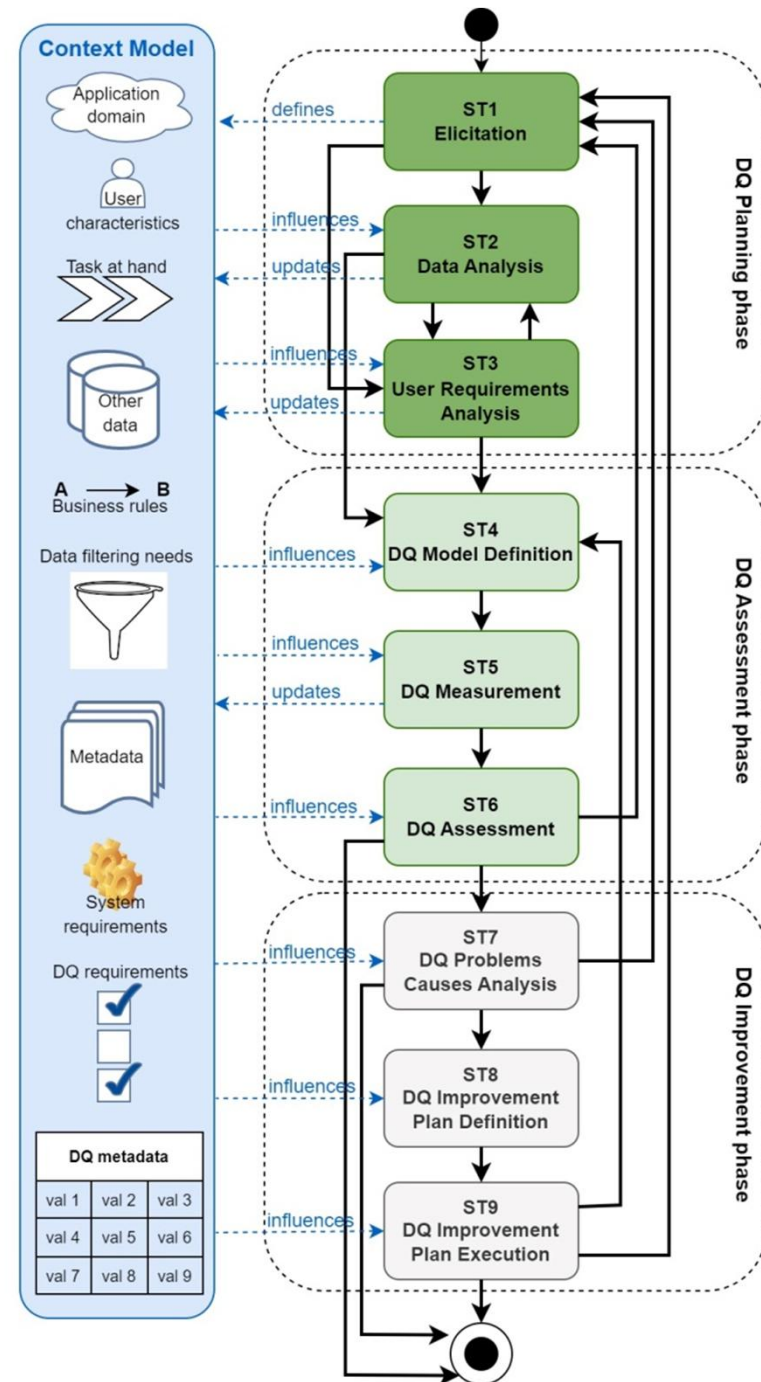
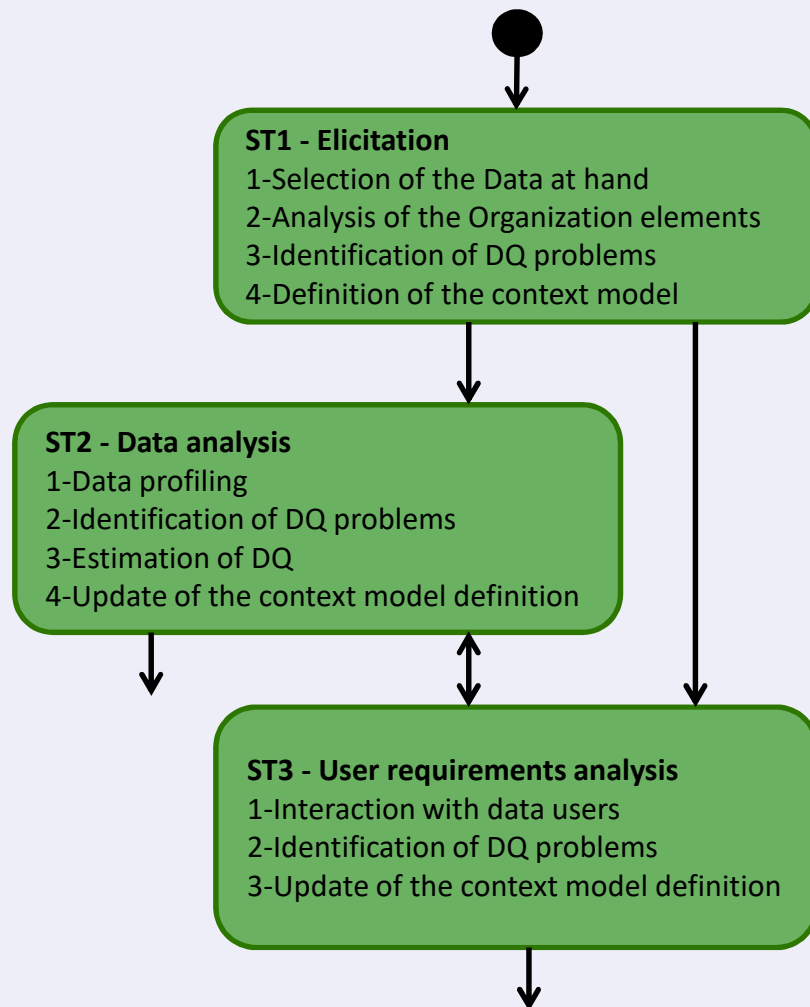
CaDQM

Context-aware Data Quality Methodology



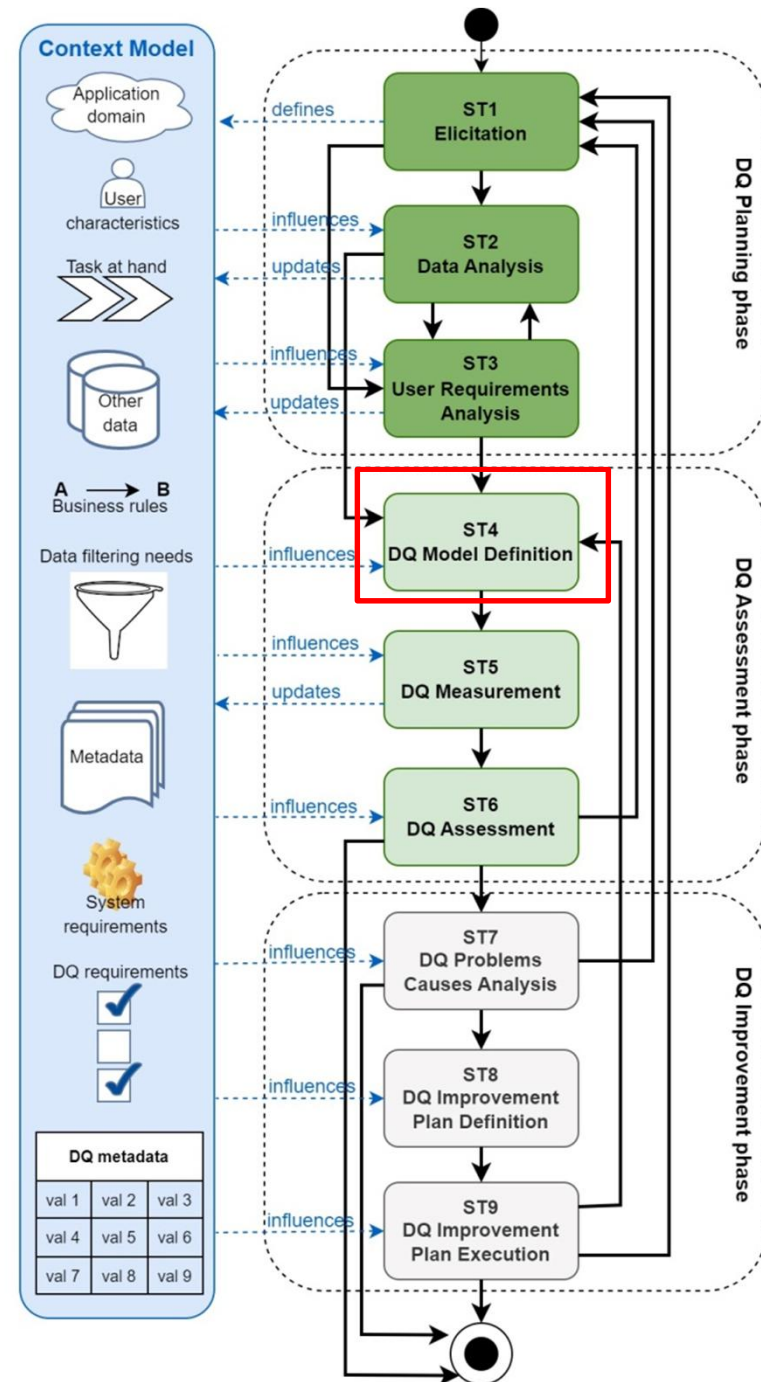
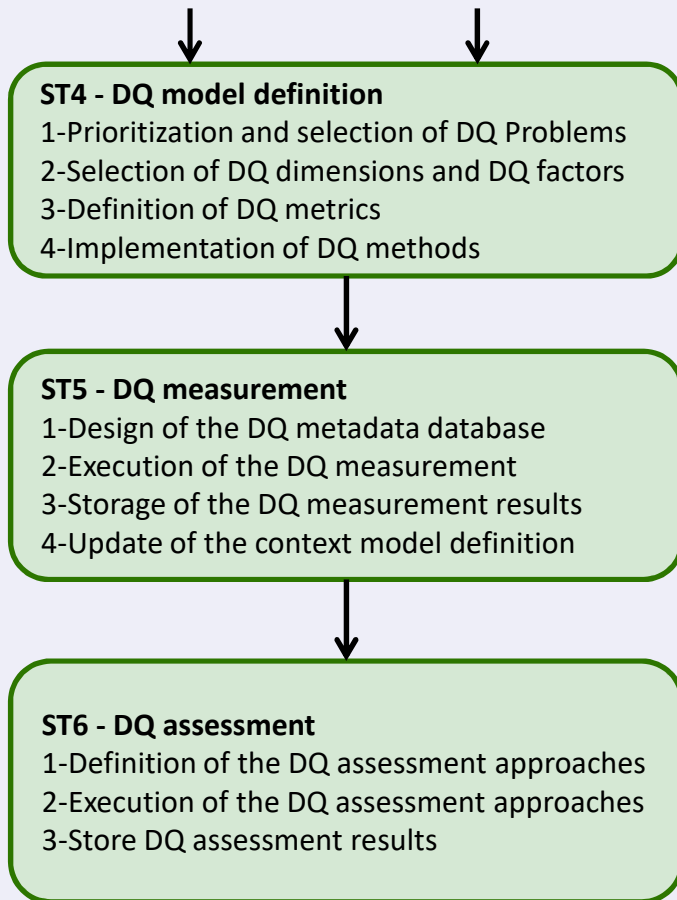
CaDQM

Phase 1: DQ Planning

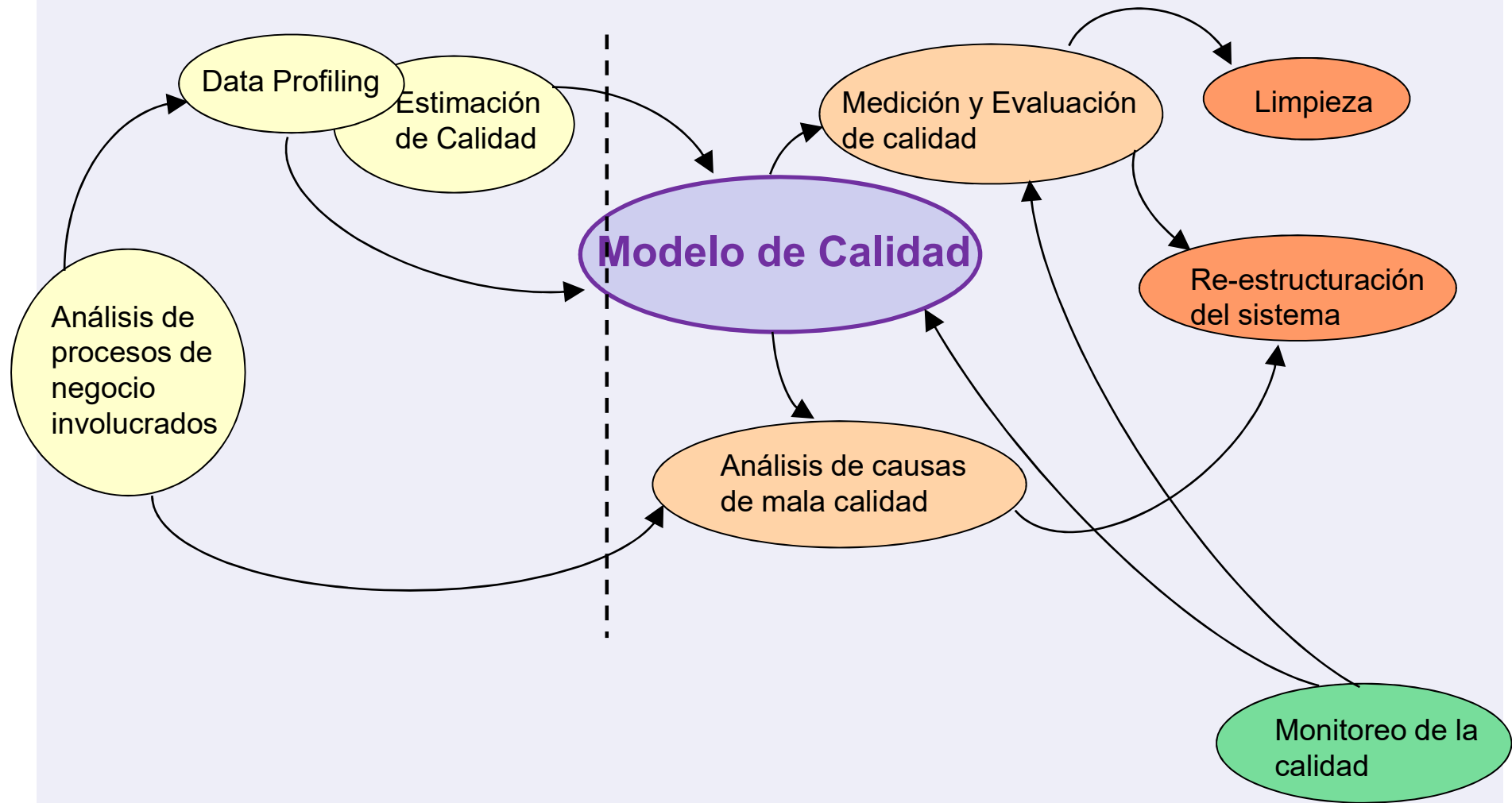


CaDQM

Phase 2: DQ Assessment



Gestión de la calidad en SI

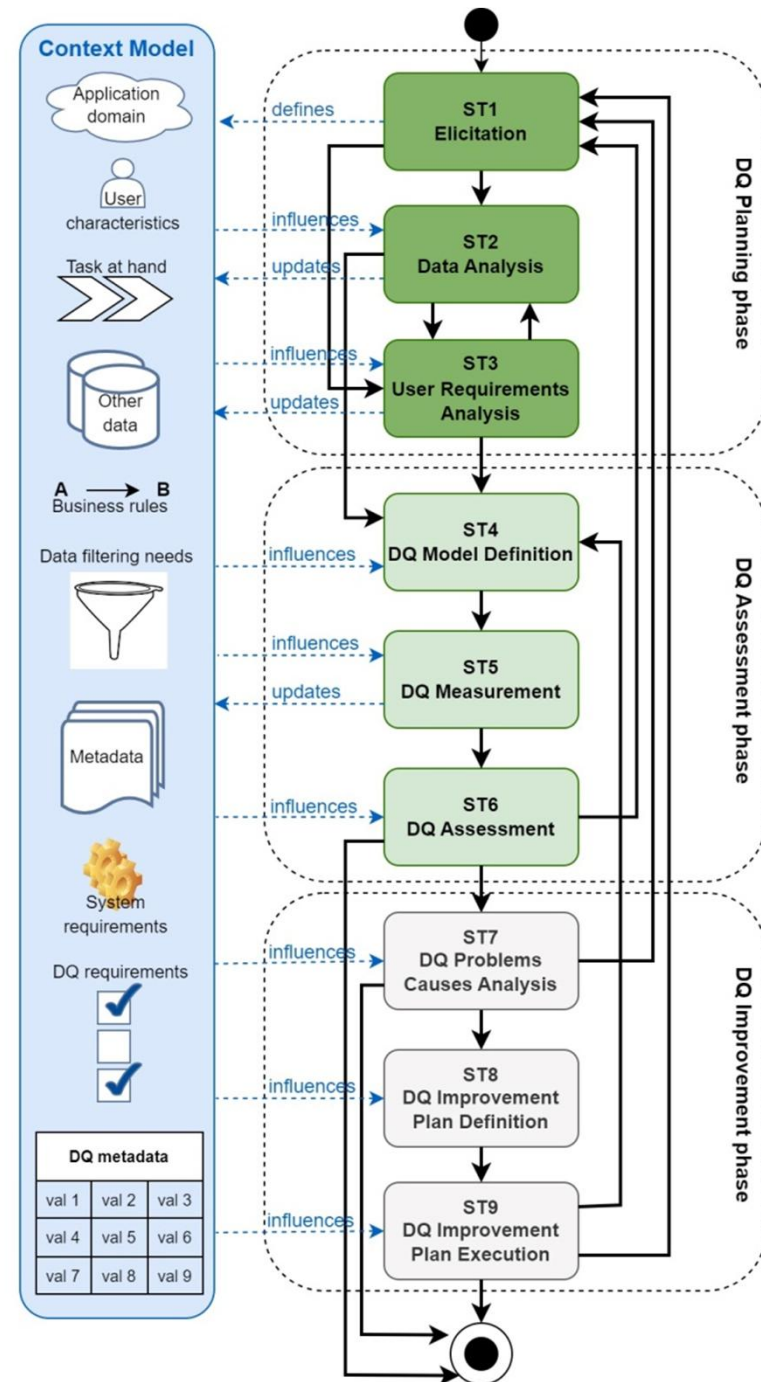
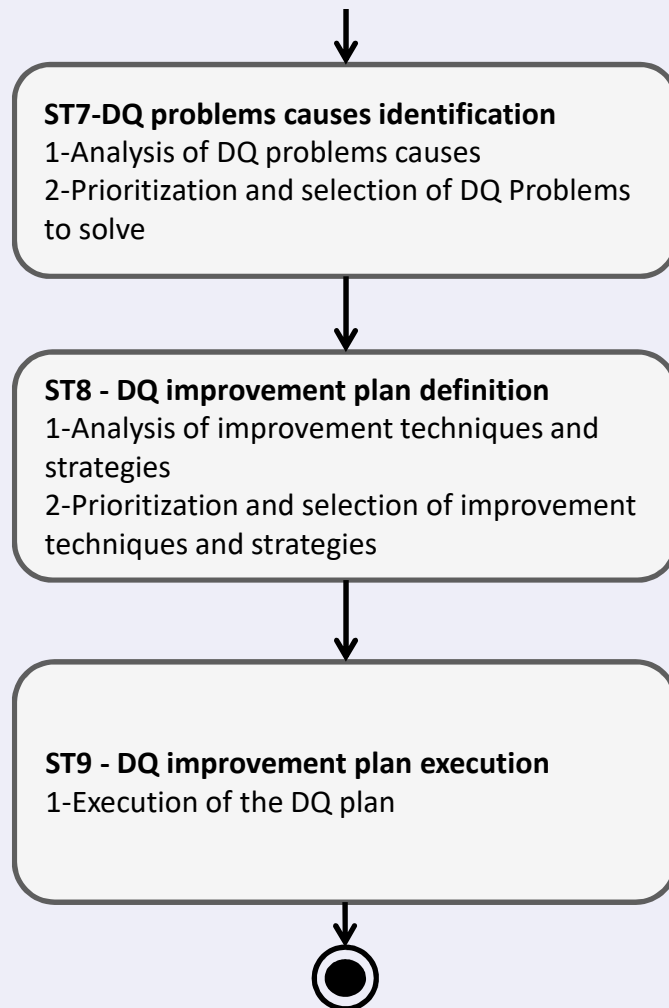


Modelo de Calidad de datos

- Define:
 - qué características de calidad se manejan
 - sobre qué datos aplican
 - cómo se miden esas características
- Para cada conjunto de datos se define un modelo de calidad particular
- Guía toda la gestión de la calidad de los datos

CaDQM

Phase 3: DQ Improvement



Bibliografía

- **Carlo Batini, Cinzia Cappiello, Chiara Francalanci, and Andrea Maurino. 2009. "Methodologies for data quality assessment and improvement". ACM Comput. Surv. 41, 3, Article 16 (July 2009), 52 pages.**
- **Carlo Batini, Monica Scannapieco. Data and Information Quality. Springer. ISBN: 978-3-319-24104-3. 2016.**
- **Corinna Cichy and Stefan Rass, "An Overview of Data Quality Frameworks," in IEEE Access, vol. 7, pp. 24634-24648, 2019.**
- **Anind K. Dey. "Understanding and Using Context". Personal Ubi Comp 5, 4–7 (2001).**
- **CaDQM: "Context-aware data quality management". Flavia Serra (2024). Tesis de Doctorado. UdelaR. PEDECIBA, Área Informática. Université de Tours.**