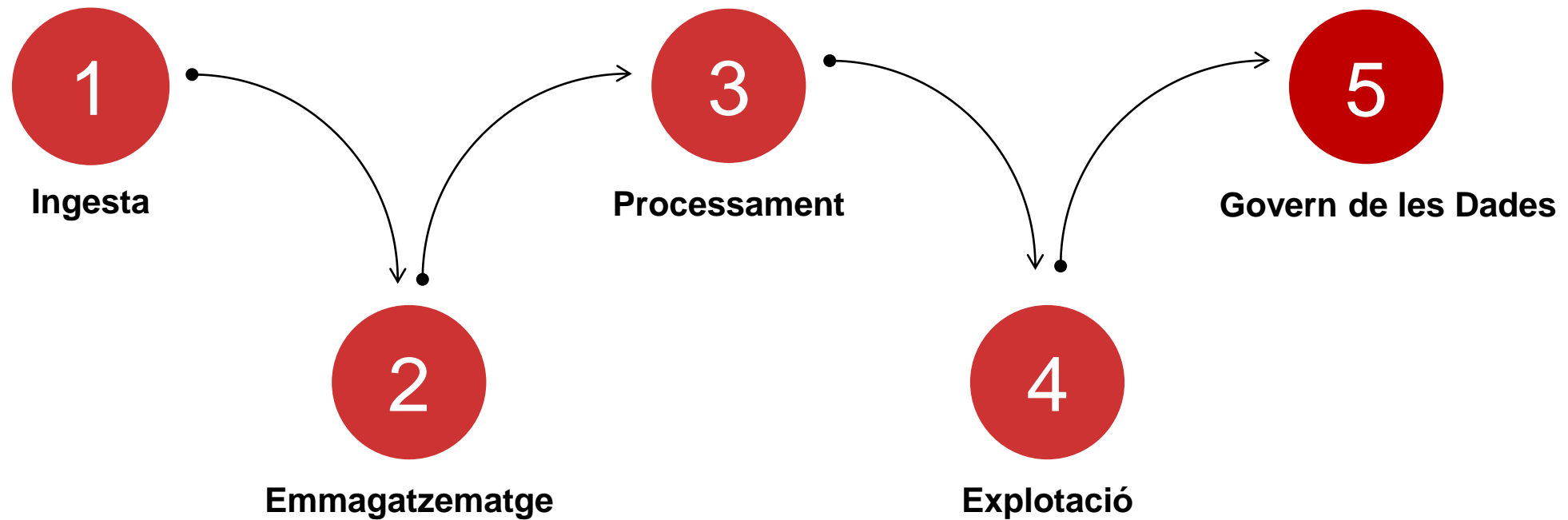


PTD – El camí de la dada

Les capacitats de la PTD

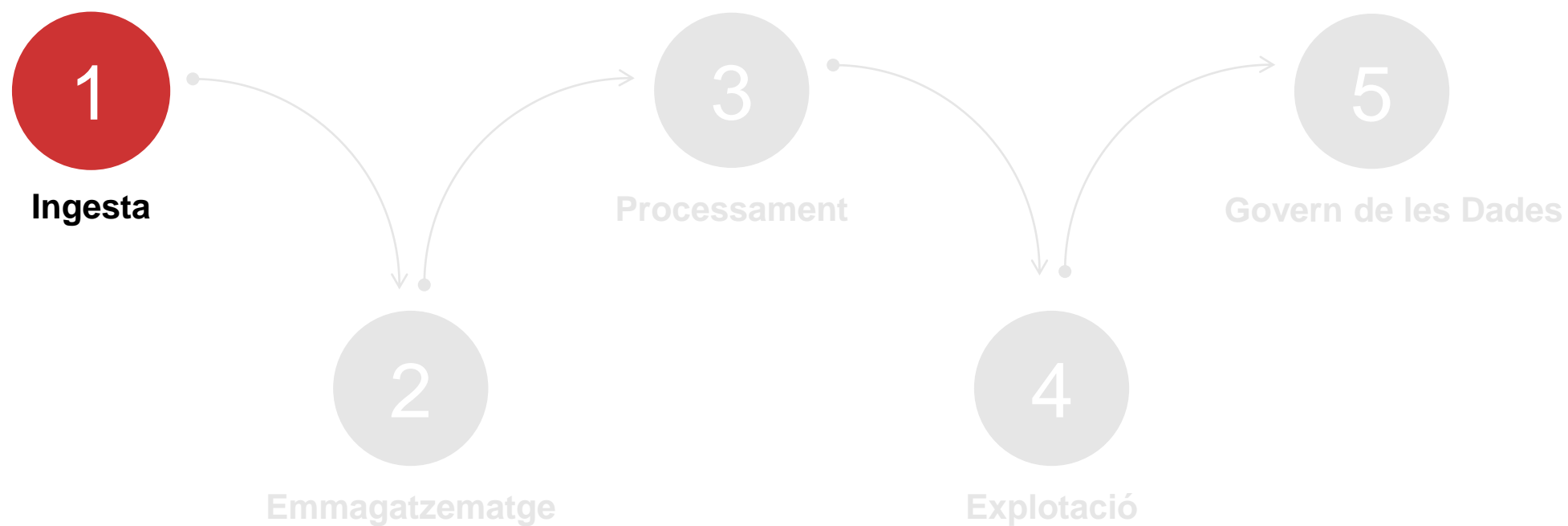
El camí de la dada - Visió a alt nivell

Com les dades viatgen i es transformen a través de la PTD



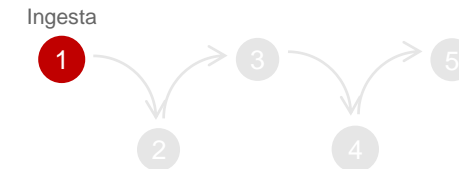
El camí de la dada - Visió a alt nivell

Com les dades viatgen i es transformen a través de la PTD



El camí de la dada - Ingesta

Què succeeix quan s'envien dades a la PTD



Comença un viatge al núvol: Les dades entren a l'entorn de la PTD



Des d'on

Les dades a ingestar poden venir de diferents orígens:

- Bases de dades d'aplicacions o sistemes analítics
- API
- Esdeveniments
- Enviament de fitxers via FTP
- Virtualització de dades d'orígens complexos



Quan

Les ingestes es podran fer amb la freqüència necessària per a cada tipus de dada i de necessitat:

- Diària
- En temps real
- Cada x temps (segons, minuts, hores, setmanes, mesos, ...)



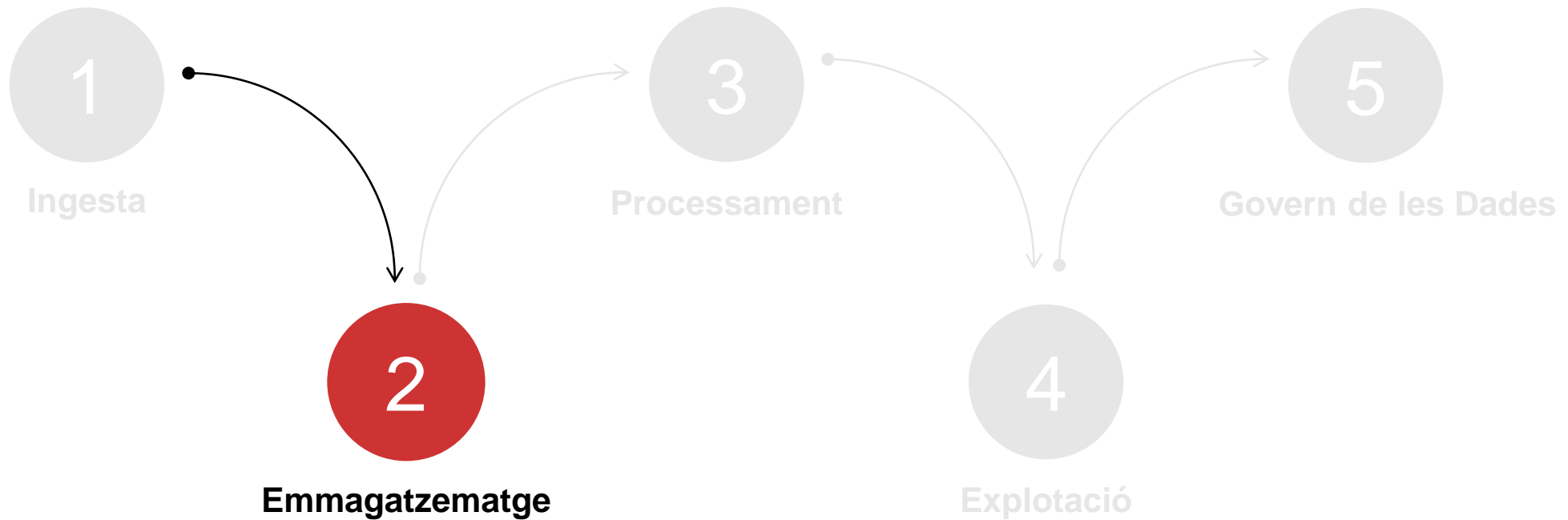
Com

Més enllà dels processos tècnics de la forma de fer les ingestes, el que sempre es complirà serà:

- Entrada en un entorn segur i robust, que compleix amb totes les directrius del CTTI.
- Les dades privades o sensibles poden ser emmascarades per protegir-les.
- Les dades estaran governades per facilitar la compartició dins de l'organització.

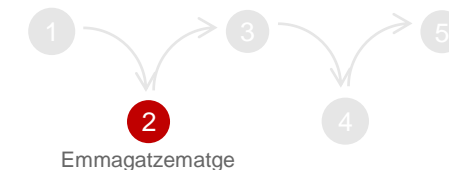
El camí de la dada - Visió a alt nivell

Com les dades viatgen i es transformen a través de la PTD



El camí de la dada - Emmagatzematge

Què succeeix quan s'envien dades a la PTD



Gran capacitat d'emmagatzematge: Qualsevol tipus de dada preparada per al seu ús



Tipus de dades

La PTD accepta qualsevol tipus de dada existent a la Generalitat. És a dir, pot emmagatzemar (i tractar / processar):

- Dades estructurades
 - Taules, bases de dades
- Dades semi-estructurades
 - JSON, XML, logs
- Dades no estructurades
 - Documents de text, imatges, vídeos, ...



Ordenació

Les dades estan ordenades i catalogades pels dominis d'informació formals de la Generalitat de Catalunya.

A més, les dades s'organitzen per capes per facilitar la seva gestió.

Hi ha un procés d'homogeneïtzació de les dades per facilitar el seu ús posterior i permetre el seu creuament.



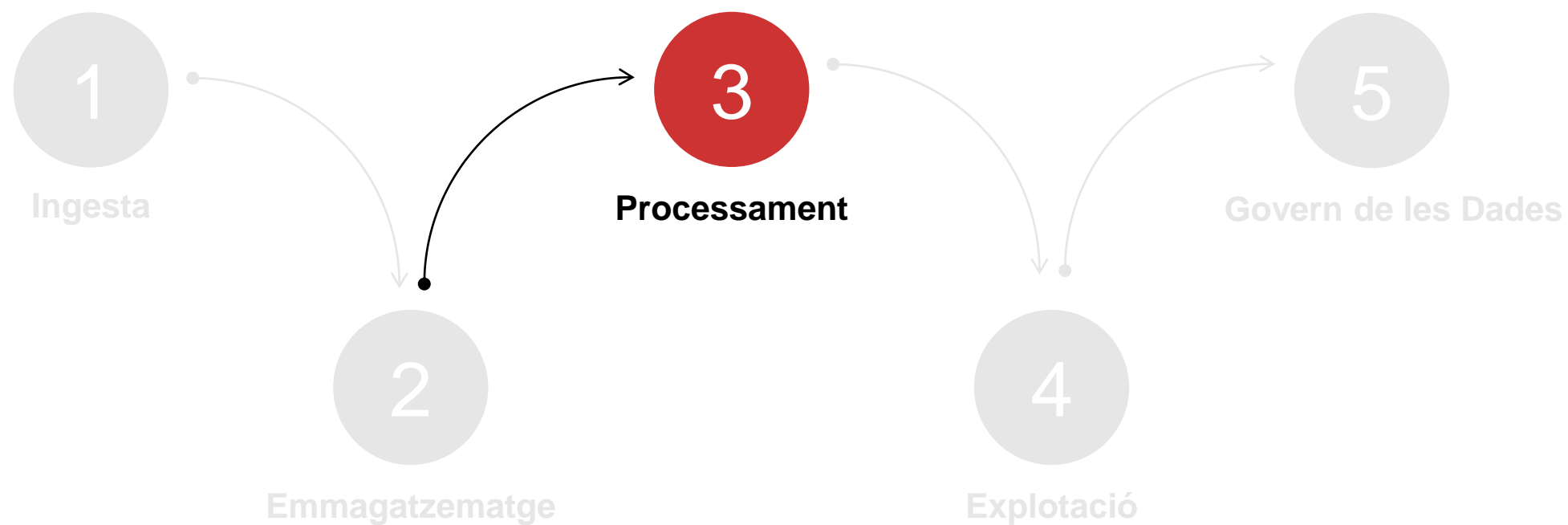
Compartir i col·laborar

Les dades a la PTD es poden compartir entre diferents departaments, sempre que ambdues parts hi estiguin d'acord i hi hagi un acord.

També hi ha dades generals a disposició de qualsevol que les vulgui incloure, de manera que es pugui millorar la funcionalitat d'algun cas d'ús.

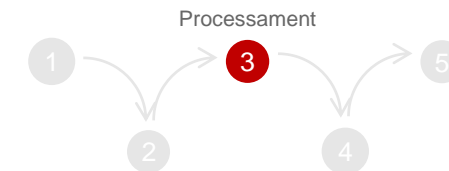
El camí de la dada - Visió a alt nivell

Com les dades viatgen i es transformen a través de la PTD



El camí de la dada - Processament

Les capacitats de processar dades a la PTD son molt àmplies



Processament sense límits: Es poden processar tot tipus de dades i volums



Processament avançat

Les eines de processament de la PTD permeten realitzar qualsevol tipus de tractament de dades tradicional, així com altres no tan convencionals (processar imatges, textos, etc.).

També disposa de capacitat d'escalar la capacitat per a tractar grans volums de dades sense interferir en altres processos.



Analítica i IA

La PTD disposa d'eines específiques per a poder desenvolupar projectes d'analítica tradicional, incloent el serveis de la STA, analítica avançada (predictiva, prescriptiva) i d'intel·ligència artificial (models d'aprenentatge automàtic i IA generativa).

També es poden crear entorns de laboratori per a científics de dades dels departaments.

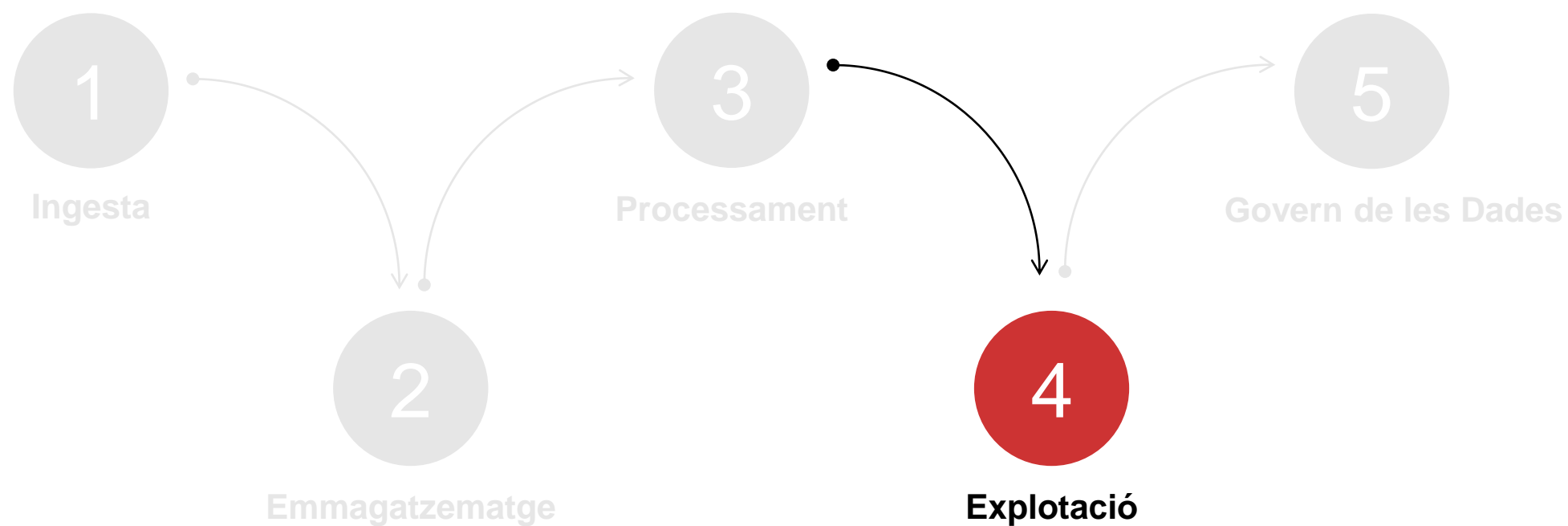


Dada única

A la PTD hi ha eines específiques per a facilitar la gestió de dades en dos grans àmbits:

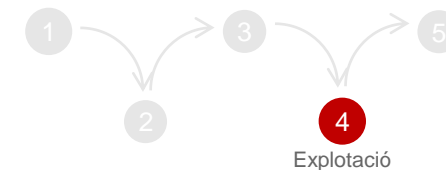
- RDM: gestió de dades de referència, per tal d'ajudar als departaments tant a crear com a fer servir dades de referència de diferents àmbits.
- MDM: gestió de dades mestres, per tal de consolidar dades de diferents orígens.

El camí de la dada - Visió a alt nivell



El camí de la dada - Explotació

La PTD permet fer un ús intensiu de les dades en diferents modalitats



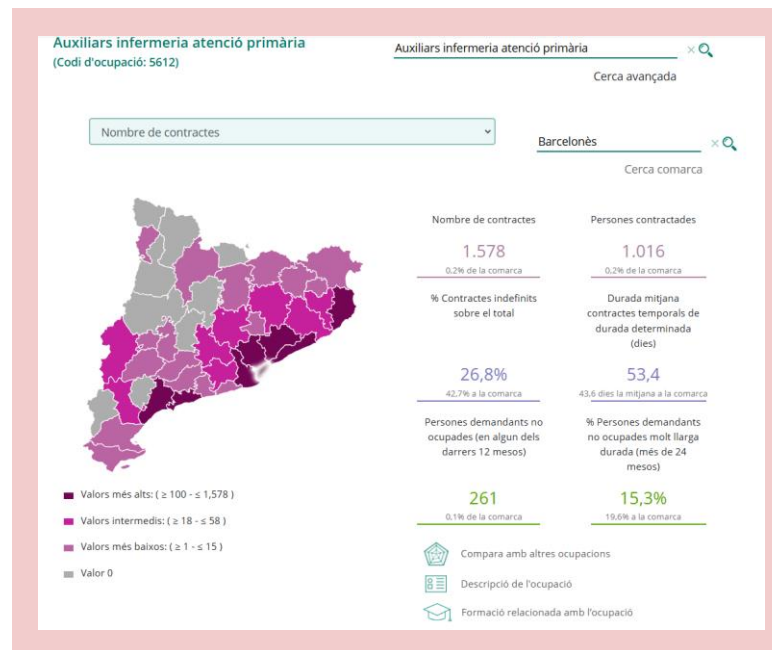
Diferents models d'explotació: Hi ha moltes maneres d'explotar les dades resultants



Explotació tradicional

Capacitat d'explotació de dades de forma tradicional:

- Eines de BI corporatives, com Power BI o Microstrategy
- Entorns d'autoservei
- Altres eines d'analítica de dades (explotació de dades geogràfiques, mineria de processos, etc...)

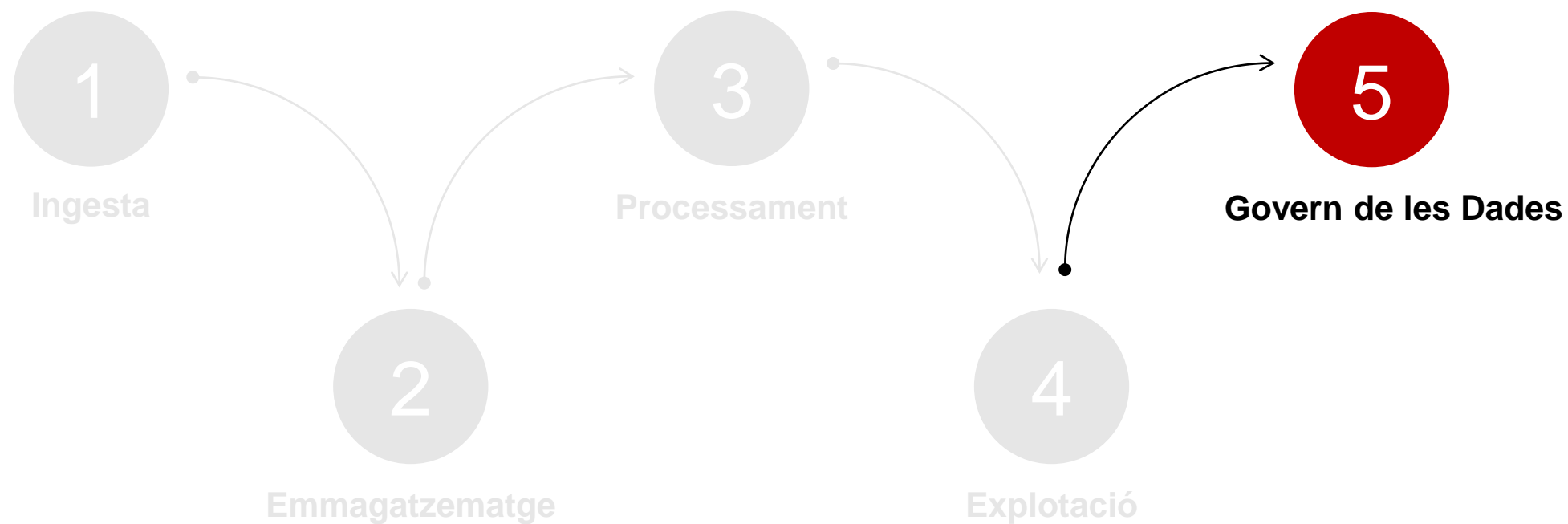


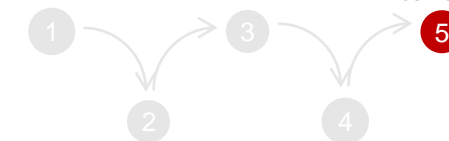
Explotació avançada

A la PTD s'hi pot realitzar:

- Ciència de dades (recerca, notebooks, ...)
- Analítica avançada i Intel·ligència Artificial
- Xatbots basats en les dades d'una unitat
- Publicació via API pel consum de tercers
- Enviament a aplicacions departamentals
- Publicació cap a dades obertes
- ...

El camí de la dada - Visió a alt nivell





El camí de la dada - Govern de les Dades

La PTD disposa d'eines i models per facilitar l'aplicació del Govern de les Dades



Dades governades i fiables: La PTD ha de millorar el coneixement i la fiabilitat de la dada



Catalogació

Les dades de la PTD estan catalogades. Això vol dir que:

- Estaran descrites en un glossari de negoci que permetrà entendre la dada
- Hi haurà un catàleg amb les dades localitzades per facilitar la seva cerca
- Tots això comprés dins d'una "botiga de dades" que permetrà buscar i localitzar dades per conceptes, glossari, descripció, etc.
- Establirà contractes de dades automàtics per poder facilitar la petició de dades i garantir el seu accés



Qualitat i llinatge

Hi ha diferents eines de qualitat que permetran dues funcions principals:

- Avaluar la qualitat de les dades (d'una taula, origen, etc.)
- Poder millorar les dades (quan sigui possible) en la seva ingesta a la PTD.

També hi ha la visió unificada de les diferents transformacions que va tenint la informació dins de la PTD (llinatge).

Això permetrà establir plans de millora de la dada des dels orígens, i avaluar com evoluciona.



Responsabilitat

Les dades a la PTD estaran sempre sota la responsabilitat del seu propietari, de forma que tant a l'hora de definir regles de qualitat, com de definir el significat de les dades, serà el responsable funcional qui ho validarà.

A més, la possibilitat d'accedir a unes dades concretes haurà de ser aprovat pel responsable de la mateixa.



Generalitat de Catalunya

Departament de Polítiques Digitals i Administració Pública

Centre de Telecomunicacions i Tecnologies de la Informació (CTTI)