

Plataforma Transversal de Dades

Estàndards tècnics

Maig 2026



1

Necessitat i avantatges

Estàndards tècnics

Necessitat i avantatges dels estàndards tècnics

Necessitat

Establir un procediment per assegurar que tots els desenvolupaments compleixen el **model de govern** i de bones pràctiques.

Solució

Capas funcionals: Múltiples capes de validacions i transformacions abans d'emmagatzemar-se per després fer una anàlisi eficient.

Nomenclatura: conjunt de guies i estàndards per definir els noms dels actius analítics.

Avantatges

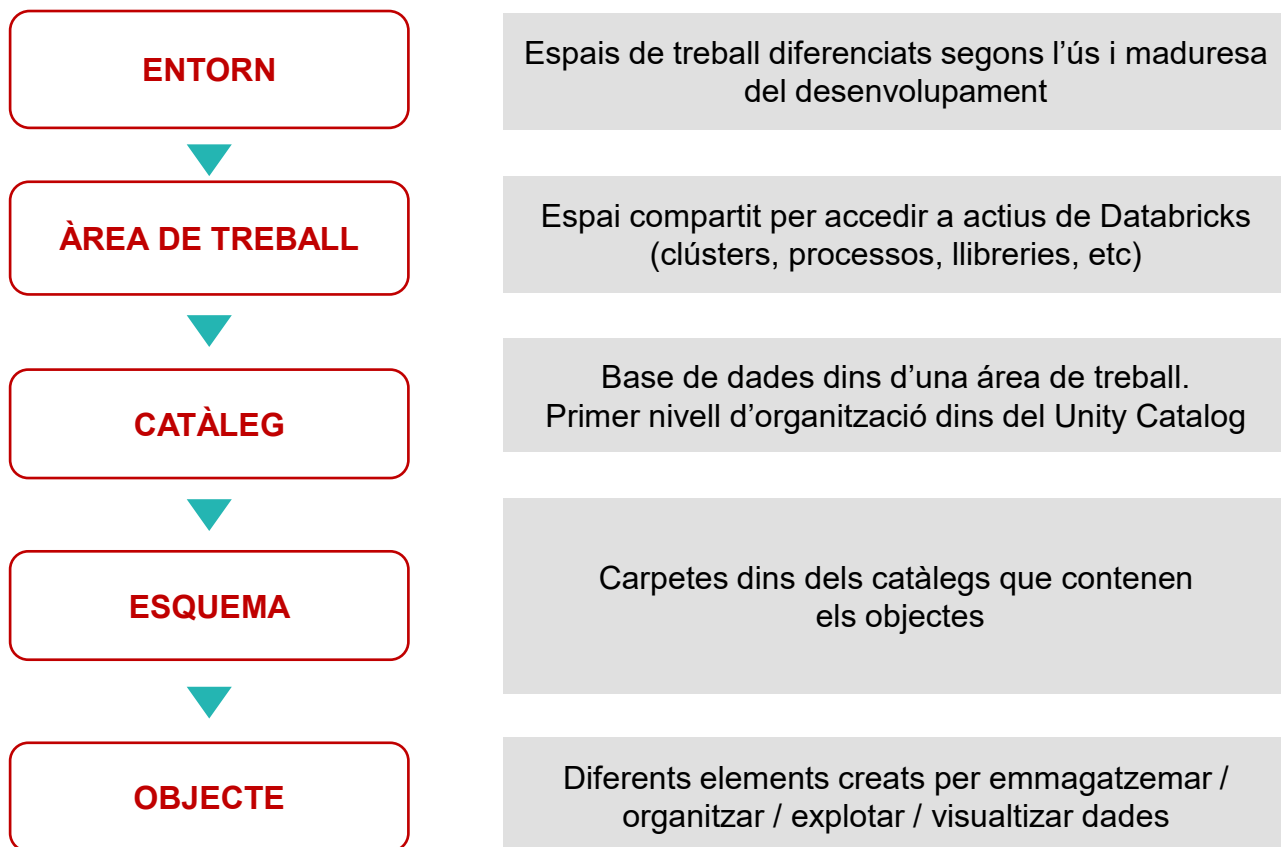
- Orientació Data Mesh
- Traçabilitat
- Auditoria
- Dades catalogades
- Consistència
- Processos reutilitzables

2

Organització de les dades a la PTD (Databricks)

Organització de les dades

Conceptes generals



3

Estructura de capes

Estructura de capes

Definició i implementació de l'estructura

Per a què serveix?

Les capes funcionals són **divisions lògiques** que distingeixen els diferents estats per on ha de passar una dada al llarg del cicle de desenvolupament i el seu processament.

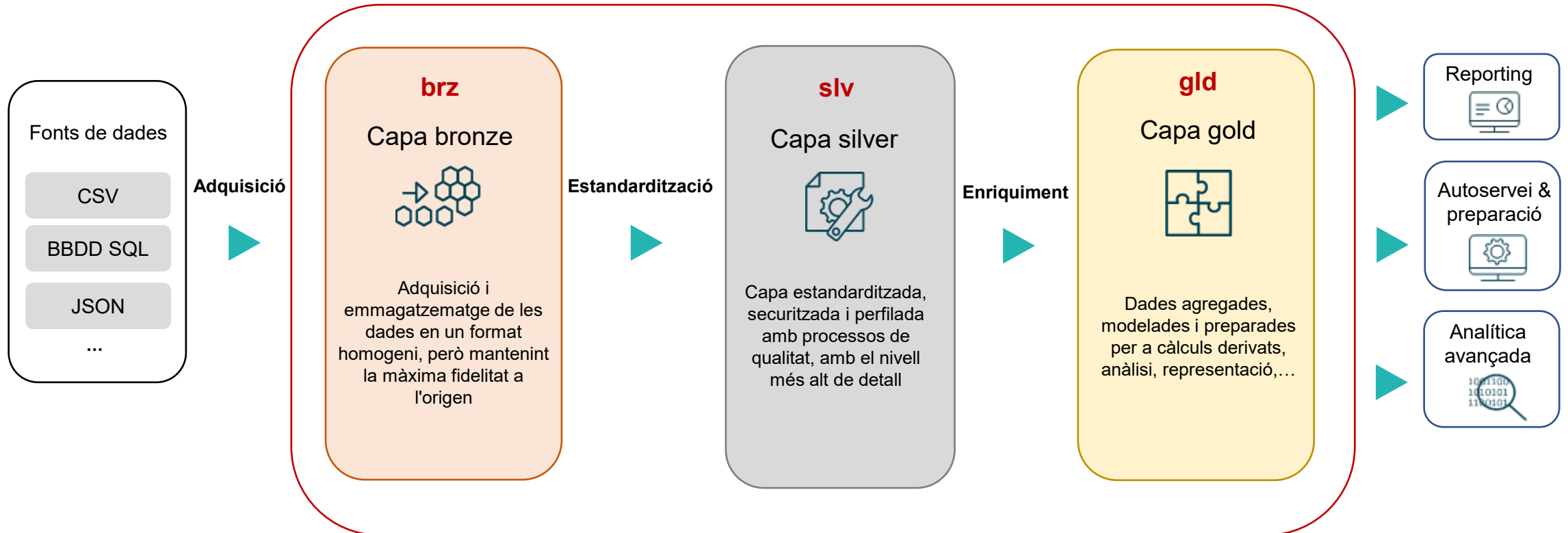
Esquema	Tipus de dada	Catalogació	Exemple
Bronze (brz)	Dada original (dada en brut, com està a l'origen)	Per aplicació origen	brz_dibica
Silver (slv)	Dada estandarditzada (noms de taula i camps normalitzats, qualitat si s'escau)	Per <u>subdomini funcional</u>	slv_museus_arxius
Gold (gld)	Dada enriquida (modificació o combinació de diverses dades estandarditzades)	Per cas d'ús	gld_sintesi

Aquesta estructura diferenciada permetrà **estandarditzar el flux de dades**, així facilitar tant una correcta definició dels **nivells de seguretat** i aplicació de permisos d'accés a la dada, com facilitar el **govern i descobriment d'aquesta**.



Ordenació funcional

Esquema de les diferents capes lògiques en la PTD



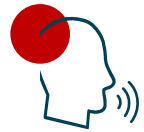
3

Estàndards tècnics

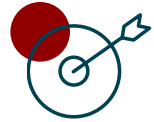
Estàndards tècnics

Resum de bones pràctiques

Per tal de poder utilitzar les dades per desenvolupar casos d'ús, és necessari que aquestes estiguin **estandarditzades** i compleixin una sèrie de requisits de nomenclatura (noms de taules i camps):



Noms **auto-explicatius**



Notacions **clares i concises**



Tot escrit **en català**



Noms de taules i camps escrits **en singular**



Noms de camps escrits **en minúscules**



Utilitzar “**_**” **com a separador** a canvi de l'espai



Utilitzar els **prefixos definits per identificar la naturalesa** de taules i camps



Evitar l'ús d'abreviatures comunes a diversos àmbits

Complir amb tots aquests requisits ens ajudarà a obtenir noms **representatius, funcionals i estandarditzats** que ens facilitaran la **comprensió dels nostres actius de dades**.

Estàndards tècnics

Polítiques i exemples aplicats

POLÍTIQUES GENERALS

Donada la naturalesa i la projecció del negoci, s'utilitza **el català** com a llengua vehicular en la notació de conceptes per taules i camps, entre d'altres.

Dins de la nomenclatura no s'utilitzaran caràcters especials com accents, dièresis o altres (com “ç” que es substituirà amb “c”)

Els **noms funcionals (de camps i taules)** han de ser **representatius** del seu contingut.

Els **camp** sempre es definiran en **singular** i les taules sempre que la semàntica ho permeti.

No existeixen restriccions en la **longitud** del nom de les taules o camps, però es recomana utilitzar conceptes simples i intuïtius (que no superin 30 caràcters).

És molt recomanable l'ús de **minúscules** (no suposa un problema tècnic ni per PTD ni per l'eina de govern, però per consistència)

Evitar l'ús d'abreviatures en camps o taules que puguin provocar confusions degut a la seva ambigüitat o el seu ús comú en d'altres àmbits.

Les diferents paraules o termes que s'utilitzin sempre es **separen** per “_”, mai amb espais.

A la capa **bronze** tots els camps han de **mantenir el seu format d'origen**. En cas que no provinquin de fitxers i necessitin adaptacions (per exemple un format de dades no considerat dins la PTD), el format d'adaptació serà tipus VARCHAR.

A totes les taules hauran d'existir **camp de control: data de càrrega** (tms_carrega) referent a la data en que la Informació va ser carregada a la base de dades; i **la data d'actulització les dades** (tms_modificació) referent a la data en què va realitzar la darrera modificació del registre.

EXEMPLE

dim_universitats

ide_ciutada en lloc de **ide_ciudadà**

ide_ciutada

ide_contracte en lloc de **ide_contractes**

num_superf_m2_utils en lloc de
num_superficie_metres_quadrats_utils

ide_vehicle

des_tipus_contracte en lloc de **des_tp_ctr**

num_superf_m2_utils em lloc de **num superf m2 utils**

turisme_intern

tms_carrega i **tms_modificacio**

Notació dels objectes

Notació de catàlegs a Databricks

En l'estructura d'emmagatzematge que existeix a Databricks, es necessària la creació d'agrupacions de bases de dades, els catàlegs. Es crearan a partir del domini i l'entorn tecnològic (**des**, **pre** i **pro**). Tots els catàlegs creats hauran de seguir la següent notació:

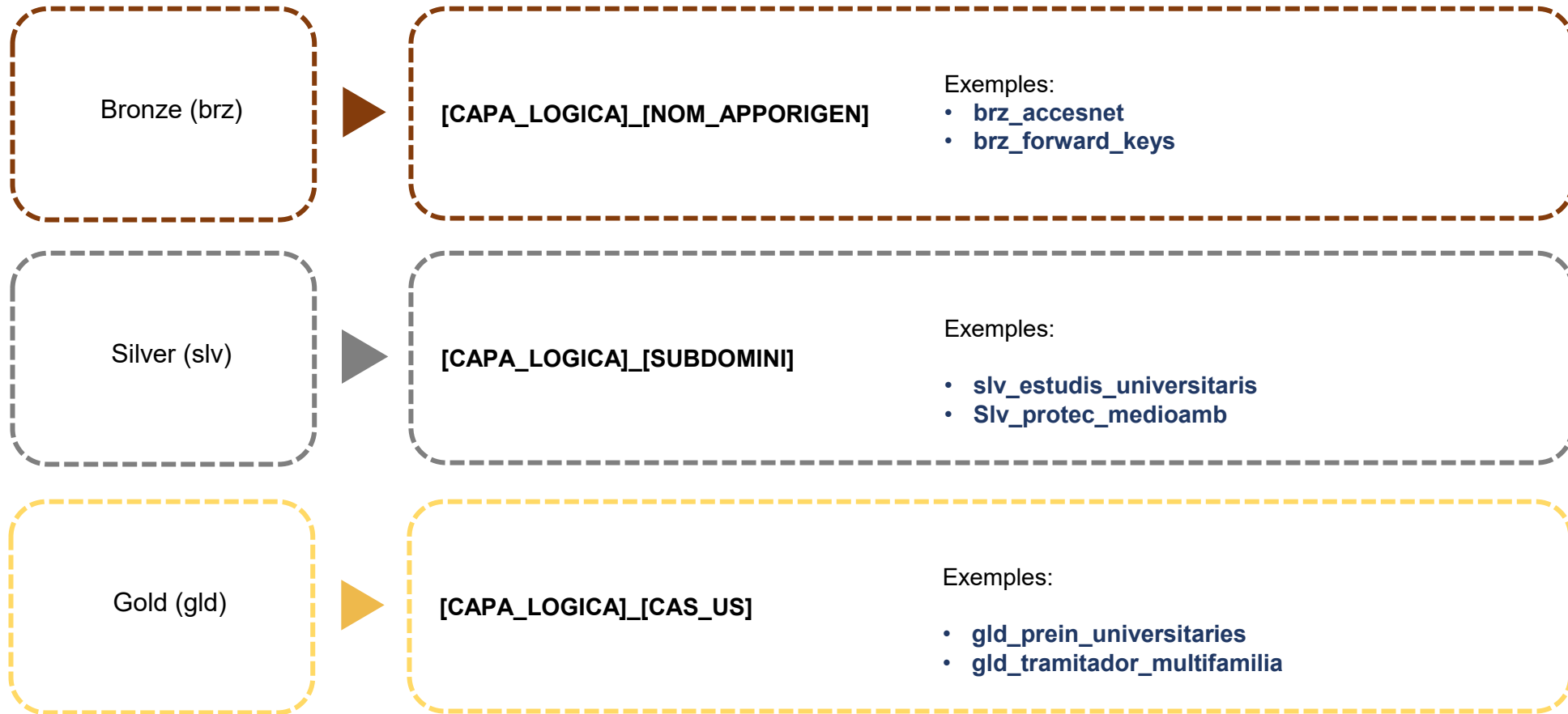
[**abreviatura domini**][_**entorn tecnològic**]; exemples: **admin_govern_pre**, **mobilitat_pro**

Codi	Acrònim	Nom del domini	Abreviatura del domini
D01	adm	Administració pública, govern i relacions institucionals	admin_govern
D02	agr	Agricultura, ramaderia, pesca i alimentació	agric_aliment
D03	cct	Comerç, consum i turisme	consum_turisme
D04	clt	Cultura, llengua i identitat	cult_llengua
D05	eco	Economia	economia
D06	edu	Educació, formació i universitats	edu_formacio
D07	ell	Esports i lleure	esport lleure
D08	fph	Finances públiques i hisenda	finances_hisenda
D09	ind	Indústria i energia	ind_energia
D10	inf	Informació i comunicació	inform_comun
D11	jud	Justícia i dret	justicia_dret
D12	mam	Medi ambient	medi_ambient
D13	mob	Mobilitat i transports	mobilitat
D14	slt	Salut	salut
D15	sem	Seguretat i emergències	segur_emerg
D16	sso	Serveis socials	serveis_socials
D17	sci	Societat i ciutadania	societat
D18	tec	Tecnologia, recerca i innovació	tecno_innova
D19	ter	Territori, infraestructures i urbanisme	territori
D20	trb	Treball	treball

Notació dels objectes

Notació de capes lògiques i volums (objectes per emmagatzament de fitxers)

Els noms d'esquemes i volums han de seguir la següent nomenclatura, **independentment** de l'entorn tecnològic (**des**, **pre** i **pro**):



Notació dels objectes

Notació d'esquema silver a Databricks

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**]; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació subdomini
D01	D01S01	Acció de govern i normativa. desplegament de l'autogovern	accio_govern
D01	D01S02	Acció exterior i relacions institucionals (CCAA, Estat i UE)	exterior_rrii
D01	D01S03	Funcionament de l'administració (règim jurídic, funció pública, procediment administratiu, contractació pública, expropiació, responsabilitat patrimonial, administració digital, etc.)	funcion_adm
D01	D01S04	Organització de l'administració i sector públic	organit_sp
D01	D01S05	Organització territorial i règim local	organit_territori
D02	D02S01	Agricultura i ramaderia	agric_ramader
D02	D02S02	Aprofitament forestal i boscos	aprof_forestal
D02	D02S03	Caça, pesca i activitats marítimes	caca_activmar
D02	D02S04	Indústria agroalimentària	ind_agroal
D02	D02S05	Sanitat vegetal i animal	sanitat_veganim
D03	D03S01	Comerç i fires	comerc_fires
D03	D03S02	Consum	consum
D03	D03S03	Defensa de la competència i de consumidors	compet_consum
D03	D03S04	Denominacions geogràfiques i de qualitat	denom_qualit
D03	D03S05	Turisme	turisme

Notació dels objectes

Notació d'esquema silver a Databricks (cont.)

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**]; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació Domini
D04	D04S01	Cultura popular i tradicional	cult_popular
D04	D04S02	Indústries culturals	ind_culturals
D04	D04S03	Llengua pròpia i símbols de Catalunya	llengua_simbols
D04	D04S04	Museus, arxius i biblioteques	museus_arxius
D04	D04S05	Patrimoni històric i artístic	patrim_histart
D05	D05S01	Caixes d'estalvis, banca, assegurances i mutualitats	banca_assegur
D05	D05S02	Cooperatives i economia social	coope_ecosocial
D05	D05S03	Empresa i sector econòmic	empresa_activeco
D05	D05S04	Mercats de valors i centres de contractació	mercats_valors
D05	D05S05	Registres de la propietat i mercantils	regist_protatmerc
D06	D06S01	Centres i serveis educatius (no universitaris)	centres_servedu
D06	D06S02	Centres i sistema universitari	sistem_universit
D06	D06S03	Ensenyament infantil, primària, secundària i batxillerat	ensenya_ipsb
D06	D06S04	Estudis universitaris	estudis_universit
D06	D06S05	Formació professional i ensenyaments de règim especial	fp_ensenyesp
D06	D06S06	Ordenació curricular i títols acadèmics	ord_curricular

Notació dels objectes

Notació d'esquema silver a Databricks (cont.)

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**]; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació Domini
D07	D07S01	Espectacles i activitats recreatives	espec_activrecre
D07	D07S02	Esport i lleure	esport_lleure
D07	D07S03	Jocs, apostes i casinos	jocs_casinos
D08	D08S01	Control financer i fiscalitat	fiscalitat
D08	D08S02	Deute públic i operacions de crèdit	deute_public
D08	D08S03	Finançament estatal i fons europeus	finanç_estatue
D08	D08S04	Hisenda pròpia i tributs	hisenda_tributs
D08	D08S05	Patrimoni	patrimoni
D08	D08S06	Pressupost i tresoreria	pressup_tresor
D08	D08S07	Subvencions i ajuts públics	subven_ajuts
D09	D09S01	Artesania	artesanía
D09	D09S02	Control metrològic i contrastació de metalls	control_metrolg
D09	D09S03	Energia i mines	energia_mines
D09	D09S04	Indústria	industria
D09	D09S05	Propietat intel·lectual i industrial	propietat_intel
D09	D09S06	Seguretat d'instal·lacions i activitats industrials	segurtat_activind

Notació dels objectes

Notació d'esquema silver a Databricks (cont.)

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**]; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació Domini
D10	D10S01	Estadística	estadistica
D10	D10S02	Informació i difusió	infor_difusio
D10	D10S03	Mitjans de comunicació i serveis audiovisuals	mitjans_comunic
D10	D10S04	Publicitat i transparència	publict_transp
D10	D10S05	Telecomunicacions i societat digital	telecom_socdigit
D11	D11S02	Entitats jurídiques	entitats_jurd
D11	D11S06	Atenció a la víctima	atencio_victima
D11	D11S07	Execució penal i reinserció	penal_reinsercio
D11	D11S01	Administració de justícia	adm_justícia
D11	D11S03	Suport Judicial	suport_judicial
D11	D11S04	Dret civil i notariat	dret_civil
D11	D11S05	Justícia restaurativa i resolució de conflictes	resolucio_confictes
D12	D12S01	Aigua i obres hidràuliques	aigua_obreshidr
D12	D12S02	Espais naturals i biodiversitat	espaisnat_biodiv
D12	D12S03	Meteorologia i canvi climàtic	canvi_climatic
D12	D12S04	Protecció del medi ambient	protec_medi amb
D12	D12S05	Residus	residus

Notació dels objectes

Notació d'esquema silver a Databricks (cont.)

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**]; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació Domini
D13	D13S01	Carreteres, ferrocarrils, ports i aeroports	ctres_ferrocaerop
D13	D13S02	Mobilitat i trànsit	mobilitat_transit
D13	D13S03	Transport de mercaderies	transp_mercader
D13	D13S04	Transport de viatgers	transp_viatgers
D14	D14S01	Assistència sanitària	assisten_sanitat
D14	D14S02	Estat de salut	estat_salut
D14	D14S03	Salut pública	salut_publica
D14	D14S04	Sistema i recursos sanitaris	sistem_sanitari
D15	D15S01	Emergències i protecció civil	emerg_proteccivil
D15	D15S02	Extinció d'incendis i salvament	incendis_salvam
D15	D15S03	Seguretat privada	seguretat_privada
D15	D15S04	Seguretat pública	seguretat_publica
D15	D15S05	Videovigilància i control de so i enregistraments	videvig_enregist
D16	D16S01	Acolliment i adopcions	acolliment_adopc
D16	D16S02	Discapacitat i dependència	discapcit_depend
D16	D16S03	Inclusió social	inclusio_social
D16	D16S04	Prestacions socials i xarxa d'equipaments	prestsoci_equipam
D16	D16S05	Protecció de menors, famílies i gent gran	protec_menfamil
D16	D16S06	Serveis assistencials i voluntariat	servassit_volunt

Notació dels objectes

Notació d'esquema silver a Databricks (cont.)

Les dades estandarditzades s'organitzen per subdomini funcional. Tots els esquemes silver hauran de seguir la següent notació:

[**abreviatura esquema silver**][**subdomini**] ; exemples: **slv_museus_arxius**, **slv_residus**

Codi domini	Codi subdomini	Nom del subdomini	Abreviació Domini
D17	D17S07	Participació ciutadana i consultes populars	particip_ciutada
D17	D17S08	Polítiques de gènere i igualtat	politiqu_igualt
D17	D17S09	Protecció de dades de caràcter personal	proteccio_dades
D17	D17S01	Afers religiosos	afers_religiosos
D17	D17S02	Atenció ciutadana i relacions amb l'administració	atenciudad_reladm
D17	D17S03	Drets humans, polítics i civils	drets_humpoltciv
D17	D17S05	Joventut	joventut
D17	D17S04	Migració	migracio
D17	D17S10	Infància i Adolescència	infancia_adolescencia
D17	D17S06	Memòria democràtica	memoria_democratica
D18	D18S01	Innovació tecnològica	innova_tecnolog
D18	D18S02	Personal investigador	personal_invest
D18	D18S03	Recerca i desenvolupament científic i tècnic	recerca_cientif
D19	D19S01	Habitatge	habitatge
D19	D19S02	Obres públiques i infraestructures	obrespubl_infrast
D19	D19S03	Ordenació del territori, paisatge i urbanisme	ordterrit_urban
D20	D20S01	Formació ocupacional	formac_ocupacion
D20	D20S02	Ocupació i inserció laboral	ocupacio_laboral
D20	D20S03	Prestacions d'atur	prestacions_atur
D20	D20S04	Prevenió de riscos laborals	riscos_laborals
D20	D20S05	Relacions laborals i negociació col·lectiva	relaborals_negoc

Notació dels objectes

Notació de taules a capes silver i gold

Per determinar el nom d'una taula de dades s'utilitzen estructures codificades per cada una de les capes lògiques, independentment a l'entorn tecnològic al que pertanyin (des, pre o pro). Totes les taules creades a les capes slv i gld, hauran de respectar la següent notació:

[**tipus taula**][**nom funcional**]; exemples: **dim_universitats**; **ref_codis_postal**; **agg_import_mensual_subvencio**

CAPA SILVER (SLV) i CAPA GOLD (GLD)

TIPUS TAULA	Determina la tipologia de la taula i tindrà una extensió fixe de 3 caràcters. A la capa slv i gld, existiran 11 tipus de taules disponibles: <ul style="list-style-type: none">• dim: Taula dimensional que serveix per a classificar o categoritzar un fet.• fac: Taula de fets que conté els valors de les mesures o indicadors de negoci.• agg: Taula agregada d'una taula de fets respecte alguna de les seves dimensions d'anàlisi.• rel: Taula relacional que s'utilitza per establir el vincle entre dos dimensions.• lkp: Taula <i>lookup</i> o diccionari que s'utilitza per a la traducció o descodificació de valors.• suk: Taula per identificar claus subrogades.• aux: Taula auxiliar, per necessitat d'un pas intermedi dins d'un procés tècnic.• ref: Taula que conté dades de referència.• mes: Taules mestres.• out: Taules de sortida que contenen els resultats finals d'un procés d'anàlisi de dades o d'aplicació d'un model d'aprenentatge automàtic.• ftr: Taules que contenen les característiques (<i>features</i>) que s'utilitzaran com a variables independents en un model d'aprenentatge automàtic.
NOM FUNCIONAL	Identifica de forma simple el concepte de negoci que descriu millor el contingut de la taula.
VISTA	Indicar si es tracta d'una vista amb la lletra "v_".

Notació dels objectes

Notació taules de caràcter tècnic

En el cas de les taules que s'utilitzen en processos tècnics i que contenen dades tècniques o de configuració que no són de negoci i són no explotables, es procedirà de la següent manera:

- No es realitzarà govern, es a dir, no és necessari seguir els estàndards tècnics de nomenclatures, ni capes.
- No es catalogaran les dades, no és necessària la presa de metadades, ni l'ús de les plantilles.
- No seran visibles pels usuaris, ja que no estaran a l'eina de govern.

[tipus taula]_[nom funcional]; exemples: **tec_mdm_ch_pos**; **ref_codis_postal**; **agg_import_mensual_subvencio**

Notació taules de caràcter tècnic

TIPUS TAULA	<ul style="list-style-type: none">• No es realitzarà govern, però es farà ús dels prefixos per identificar-les:<ul style="list-style-type: none">• tec: Taula tècnica, de configuració interna de la solució.• log: Taules de logs o de auditoria i registre de canvis o històrics.• tmp: Taula temporal utilitzada com a pas intermedi per als desenvolupaments.
NOM FUNCIONAL	Identifica de forma simple el concepte de negoci que descriu millor el contingut de la taula.

Notació dels objectes

Notació dels camps

El nom dels camps d'una taula és un element vital per a poder identificar la informació de forma coherent. Per tant, s'han definit les següents tipologies de camps per tal de facilitar la comprensió dels conceptes. Cada tipus de camp ha d'anar associat a un format de dades predefinit que garanteixi la integritat i homogeneïtat de la informació. Tots els camps de SLV i GLD han de complir aquesta notació.

[tipus_camp]_[nom_camp]

TIPUS CAMP	DESCRIPCIÓ TIPUS DE CAMP	FORMAT DATA LAKE	FORMAT QUERY ENGINE	EXEMPLE
ide	Identificador. Es recomana utilitzar identificadors del tipus integer excepte en casos on pugui ser necessari tipus varchar.	INT o NVARCHAR	INT, INTEGER, BIGINT, SMALLINT, TINYINT, BYTEINT // VARCHAR, CHAR	ide_port
dat	Data (<i>date</i>). Per defecte tindrà una precisió a nivell dia.	DATE	DATE	dat_inici_contracte
hrs	Data i hora. Per defecte tindrà una precisió nivell dia i hora.	DATETIME	DATETIME	hrs_venciment_vigencia
tms	<i>Timestamp</i> .	TIMESTAMP	TIMESTAMP	tms_carrega
des	Descriptor. Els descriptors podran tenir associat el sufix <code>_{idioma}</code> per a determinar l'idioma del seu contingut que tindrà. Aquest sufix tindrà una extensió fixe de 2 caràcters (ISO 639-1).	NVARCHAR	VARCHAR, CHAR	des_tipus_contracte
flg	<i>Flag</i> o Booleà (boolean).	INT o NVARCHAR	BOOLEAN	flg_vigencia
imp	Import. Els imports podran tenir associats el sufix <code>_{tipus moneda}</code> on s'especifica si el valor de la moneda es local ("LC") o l'identificador propi de la moneda d'extensió fixe 3 caràcters (ISO 4217).	DECIMAL	DECIMAL	imp_venta
num	Indicador numèric sempre i quant no sigui un import o percentatge.	DECIMAL	DECIMAL	num_superf_m2_utils
per	Indicador numèric per percentatges.	DECIMAL	DECIMAL	per_aprovat
url	Utilitzat per anomenar camps que contenen una direcció web (URL).	NVARCHAR	VARCHAR, CHAR	url_tripadvisor

Notació dels objectes

Notació dels camps (cont.)

TIPUS CAMP	DESCRIPCIÓ TIPUS DE CAMP	FORMAT DATA LAKE	FORMAT QUERY ENGINE	EXEMPLE
obj	Objecte. Tipus de dada complexa composta de dades amidades	OBJECTE (Mongo db)	STRUCT<...> o STRING	obj_alumne
arr	Array. Col·lecció d'elements del mateix tipus	ARRAY (Mongo db - databricks)	ARRAY <element_type>	arr_cursos_matriculas
map	Map. Diccionari de dades clau-valor	MAP (Data lake - Databricks)	MAP <key_type, value_type>	map_codis_licencia
stu	Struct. Tipologia "objecte" equivalent en Databricks	STRUCT (Data lake - Databricks)	STRUCT <...>	stu_alumne
img	Imatge. camp destinat a emmagatzemar o referenciar un arxiu gràfic	BINARY (Data lake - Databricks)	BINARY	img_perfil
doc	Document. Camp que emmagatzema o referencia un fitxer documental associat	BINARY (Data lake - Databricks)	BINARY (o STRING si es referencia/URL)	doc_factura



Generalitat de Catalunya

www.gencat.cat