

# **1. MODULUA: INFORMAZIOAREN ANTOLAKETA ETA EGITURA**

*1 gaia: Datu baseen kontzeptua eta Garrantzia*

**Leire Aldaz, Begoña Eguía eta Leire Urcola**



## ***1. MODULUA: Informazioaren Antolaketa eta Egitura*** ***1 Gaia: Datu baseen kontzeptua eta garrantzia.***

### **Aurkibidea.**

- Sarrera.
- Datu baseak vs fitxategiak.
- Datu-base sistema baten ezaugarriak/abantailak
- Datu base sistemen abstrakzio mailak.
- Datu-baseak vs DBKS.

### **Sarrera.**

Ukaezina da historian zehar informazioa jaso ahal izateak izan duen garrantzia. Era berean, ezinbestekoa da informazioa ondo antolatzea informazio hori kudeatuko dutenen lana erraztu nahi bada. Demagun, adibidez, banketxeak sortu ziren garaia. Sasoi hartan liburuetan idazten zituzten diruen mugimenduak, informazioa gordetzeko era bakarra zelako. Informazio hori ezin zen jatorrizko kokagunetik kanpo irakurri, ezin ziren bezero jakin baten mugimenduak zerrendatu, etab. Gaur egun inolako arazorik gabe egiten dira horrelako eragiketak.

Bestalde, informazioaren bolumena izugarri hazi da. Bolumen hori kudeatu beharrak datuak biltzeko eta kudeatzeko sistemak erabat hobetsi ditu. Datuak biltzeko sistemak izugarri hobetu ziren 60ko hamarkadan, datu-baseen sorrerarekin alegia. Sistema horiek datuak era osoan eta egonkorrean kudeatzen dituzte Datu Base Kudea Sistema deritzon aplikazioari esker.

### **Datu-baseak vs fitxategiak.**

Datu-baseen aitzindari izan ziren fitxategiak eta gaur egun oraindik zenbait kasutan erabili egiten dira fitxategietan oinarritutako aplikazioak.

Fitxategiek denbora luzerako eta kopuru handian pilatzen dituzte datuak. Baina ezbeharren bat gertatuz gero, informazioa galtzea ohikoa izaten da, eta berreskuratzeko era bakarra segurtasun kopiak egitean datza. Gainera, segurtasun kopiak egiteko erantzukizuna bete-betean dagokio erabiltzaileari, bera izango baita egunero kopia horiek egin beharko dituen. Hala era, ustekabeak datuak galduz gero, segurtasun-kopiatik berreskuratu beharko dira. Baina azken kopiatik ezbeharra izan arte gindako lana galdu egiten da. Datu-baseekin, ordea, alferrik galdutako lana erabat minimizatzen da.

Horretaz aparte, fitxategi-aplikazioetan ez da batere eroso gertatzen bertako datuak kontsultatzeko eta manipulatzeko modua. Gainera, fitxategietan oinarritutako

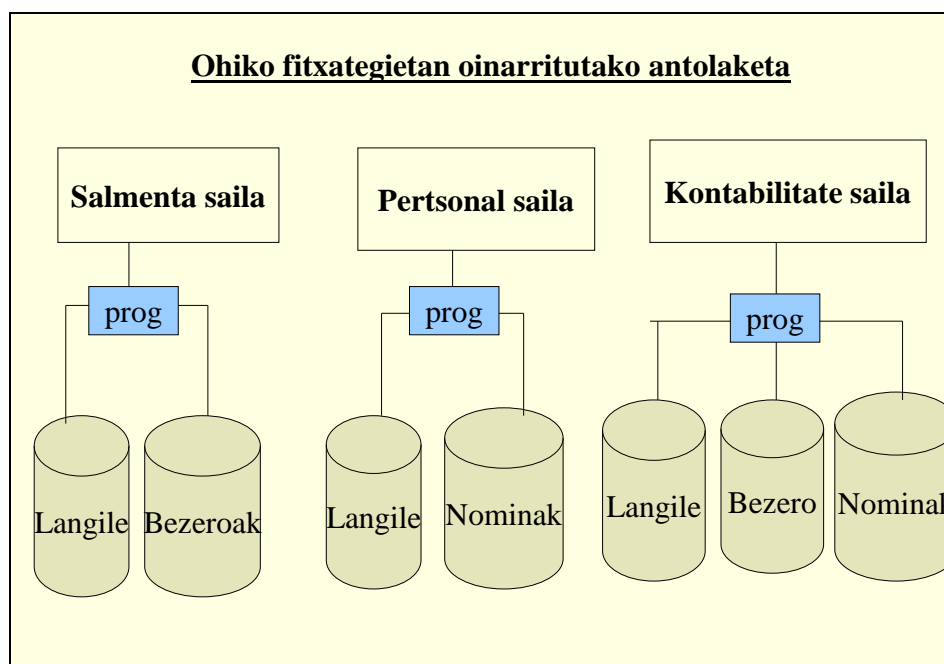


## 1. MODULUA: Informazioaren Antolaketa eta Egitura

### 1 Gaia: Datu baseen kontzeptua eta garrantzia.

aplikazioek ez dituzte aurreikusi normalean hainbat erabiltzailek aldi berean egindako atzipenak.

Fitxategietan oinarritutako aplikazioei, **gertaerei zuzendutako sistemak** deritze, egin beharko dutena analizatu eta gero aplikazioak beharrezkoak dituen fitxategiak sortzen baititu. Ondoren, beste aplikazioen bat behar izanez gero, horren analisia egin eta beharrezkoak dituen fitxategiak sortzen dira; hau da, fitxategiak aplikazio bakoitzaren mendekoak dira eta sortutako fitxategiak, elkarren artean guztiz independenteak. Hori dela eta, aplikazio kopurua hazten doan heinean, fitxategiak ugaritu egiten dira. Gainera, sarritan datuak errepikatuta agertzen dira hainbat fitxategitan, ordenagailuaren memoria alferrik betez. Eta, are txarragoa dena: datuen sendotasunik ezak zentzugabeko informazio pilaketa eragin lezake, datu berberak hainbat fitxategitan zenbait balio ezberdinekin errepikatuta agertu baitaitezke.



Adibidea:

Demagun eskola bateko idazkaritzan ikasleak kudeatzeko aplikazioa ezarri beharra dagoela eta beste aplikazio bat irakasleen gelan, irakasleek ikasleen notak kudeatzeko.

Idazkaritzako aplikazioak ikasleak matrikulatu, ezabatu eta ikasle guztien zerrendak atera beharko ditu. Horretarako, IKASLEAK izeneko fitxategia erabiliko du. Bere egitura ondorengo datuz osatuta egon liteke: Matrikula zenbakia, ikaslearen Nortasun Agiria, izena, deiturak, helbidea, ...



## *1. MODULUA: Informazioaren Antolaketa eta Egitura* *1 Gaia: Datu baseen kontzeptua eta garrantzia.*

Bestetik, irakasleen gelan, notak sartzeko aplikazioarekin ikasleen notak sartu eta kase bateko ikasleen noten zerrenda egin nahi da. Horretarako, besteak beste, TALDEAK eta NOTAK izeneko fitxategiak erabiliko dira, ondorengo datuekin:

TALDEAK: Ikaslearen Nortasun Agiria, izena, deiturak, helbidea, taldea, etab.

NOTAK: Ikasgaia-1, Ikasgaia-2, Ikasgaia-3, ikaslearen izena, deiturak, helbidea, taldea, etab.

Aipatutako adibidean, idazkaritzako aplikazioaren bidez ikasleren bat ezabatzen bada, eragiketa hori ez da TALDEAK fitxategian islatuko, bigarren aplikazio horrek ez baitu fitxategi hori kudeatzen. Hori dela eta, fitxategian ez dagoen ikasle baten notak kudeatzen aritzea gerta liteke. Ondorioz, prozesuei begira garatutako aplikazioetan datuen programarekiko daukaten mendekotasuna nabaritzen da alde batetik, eta malgutasun falta bestetik. Gainera, edozein helbide aldaketak adibidez, arazoak ekar litzake.

Eragozpen horiek direla eta, **datuei zuzendutako sistemak** garrantzi handia hartu dute. Horien artean koka daitezke datu-baseen sistemak.

Hirurogeiko hamarkadan, informazioaren egitura datu-base batean deskribatzeko aukera ematen zuten ereduak agertzen hasi ziren. Horrela, aplikazioen eta datuen pilaketaren arteko independentzia handiagoa lortu zen. Izan ere, elkarrekin erlazionatutako datuen bilduma baino ez da datu-basea. Bilduma horrek mundu errealaren zati bat (problema) adieraziko du. Adierazi nahi den mundu errealaren zatia bada ere, pertsona bakoitzak bere erara ikusiko du, era subjektiboan, eta era jakin batean adieraziko du, unean uneko komunikazio-tresna egokiena erabiliz.

### **Datu base sistema baten ezaugarriak/abantailak**

Fitxategi sistemekin konparatuz, datu baseen sistemek abantaila asko eskaintzen dituzte.

- **Datuen sendotasuna eta erredundantziaren kontrola:** fitxategi sistemetan datu bera hainbat fitxategitan errepikaturik egon daiteke, eta horrek lekua alferrik galtzea eta datuen sendotasunik eza eragiteko arriskua dakar. Printzipioz, datu baseen sistemetan fitxategi guztiak daude integratuta, eta horregatik, ez dago datuak errepikatzerik. Dena den, alde aurretik aipatu den bezala, batzuetan, datuen arteko erlazioak errazago irudikatzeko edo prestakuntzak hobetzeko, erredundantzia minimoa mantentzen da. Hori bai, datua behin baino gordetzen ez bada, behin baino ez da eguneratu behar



## *1. MODULUA: Informazioaren Antolaketa eta Egitura*

### *1 Gaia: Datu baseen kontzeptua eta garrantzia.*

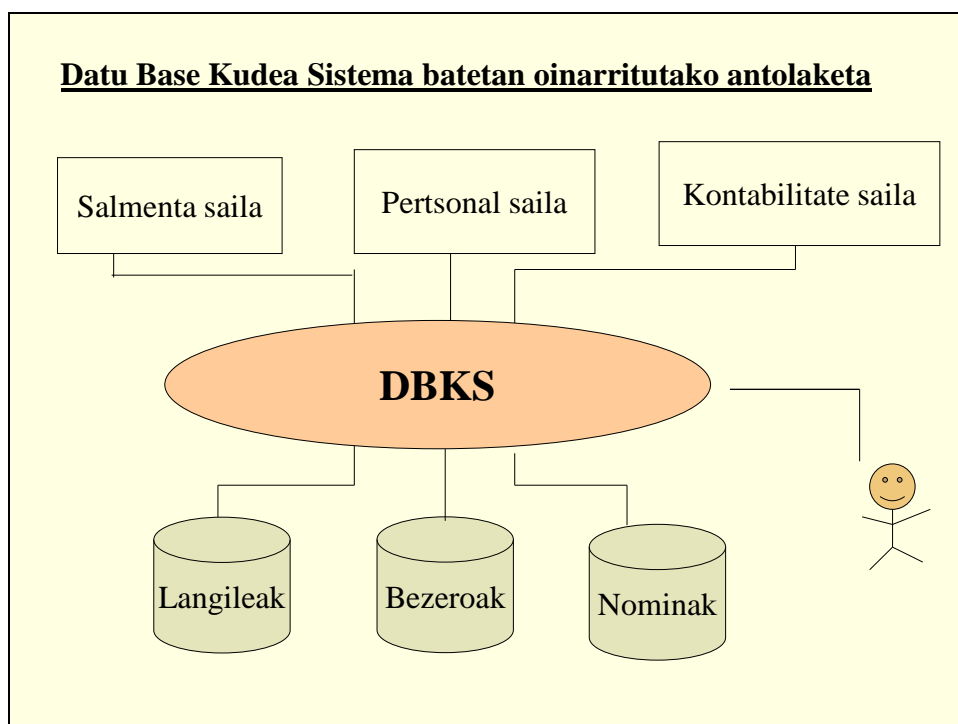
izango, eta automatikoki eskuragarri egongo da erabiltzaile guztientzat. Datua errepikatua badago, baina sistemak hori jakin badaki, sistemak berak eguneratuko ditu beste kopia guztiak.

- **Segurtasuna eta datuak konpartitzea:** fitxategi sistemetan fitxategia erabiltzen duena da fitxategiaren jabea. Datu baseetan datuak erabiltzaile guztien eskura daude, baina erabiltzaile bakoitzak bere baimenen arabera atzituiko ditu. Gainera, sorturiko aplikazio berriek dagoeneko existitzen diren datuak erabili ditzakete, eta segurtasun kopien eta berreskuratze zerbitzuen hobekuntza ere badago, fitxategi-sistemetan ez bezala. Azken horietan erabiltzailearen ardura da segurtasun kopiak egitea. DBKsek berriz, minimora gutxitzen dute alferrik galtutako lana. Eta segurtasun kopiak egiteko tresnak eskaintzen dituzte.
- **Datuak integratuta dauden heinean errazagoa da estandarrak betetzea datuen formatuan,** dokumentazioan, atzipen arauetan eta eguneratze prozeduretan. Gainera, datuen independentziari esker, mantentze lana errazago egiten da. Fitxategi sistemetan, fitxategien egitura programetan deskribatzen denez, datuen egituran edo datuak pilatzeko eran aldaketak gertatzen direnean, programak aldatu egin behar dira. Baina DBKS-etan ez da halakorik gertatzen, datuen eta aplikazioen deskripzioak banaturik daudelako. Banaketa horri independentzia deritzo eta bi mailatan gauzatzen da: batetik, independentzia fisikoak bermatzen du datuak biltegitratuta dauden erak egitura logikoan eraginik ez izatea. Hau da, nahiz eta biltegitratze fisikoan aldaketak egon, datu horietara heldu behar duen erabiltzaileak ez du programa aldatu behar izango. Eta bestetik, independentzia logikoa dago: datu baseari elementuak gehitzeak, kentzeak edo aldatzeak datu basea kudeatzen duten programetan eraginik ez izatea ahalbideratzen du.
- Datuen atzipen hoberako, **DBKsek galdeketa egiteko lengoaiak** eskaintzen dituzte. Lengoaia horiek erabiliz gero, aplikazioak programatzea ez da beharrezkoa.
- DBKsek datu berebereri aldi berean atzipen ugari egitea baimentzen dute; hots, konkurrentziarik badago, bera arduratuko da dena kontrolatzeaz, datuen osotasuna eta sendotasuna uneoro zaintzeko.



## 1. MODULUA: Informazioaren Antolaketa eta Egitura

### 1 Gaia: Datu baseen kontzeptua eta garrantzia.



Desabantailak ere badituzte, ordea:

- **Zailtasuna eta tamaina:** DBKS-ak programa multzo handiak eta konplexuak dira eta, etekin ona lortzeko, ondo ulertu behar dira. Horrez gain, programa eta datu multzo handi eta konplexu horrek leku asko behar du, bai diskoan, bai memorian.
- **Arlo ekonomikoa:** Batzuetan fitxategi sistema datu baseen sistema batekin ordeztzea oso garestia da; aplikazioa aldatzeaz gain datu baseen sistema ezarri behar da, langileei prestakuntza berezia eman behar zaie, etab. Gainera, ordenagailuek baliabide egokidunak eta pilotzeko espazio handikoak izan behar dute, DBKS-ak ondo funtziona dezan.
- **Prestazioak:** normalean fitxategi sistemak aplikazio jakin baterako sortzen dira eta, beraz, oso prestazio onak eskaintzen dituzte. DBKS-ak, berriz, aplikazio ugari erabiltzen dituzte eta, horregatik, aplikazio batzuk lehen baino astiroago joan daitezke.

Dena den, abantailak gehiago dira; horrela ez balitz, datu baseak egun dauden tokira ez ziren inoiz heldu izango.

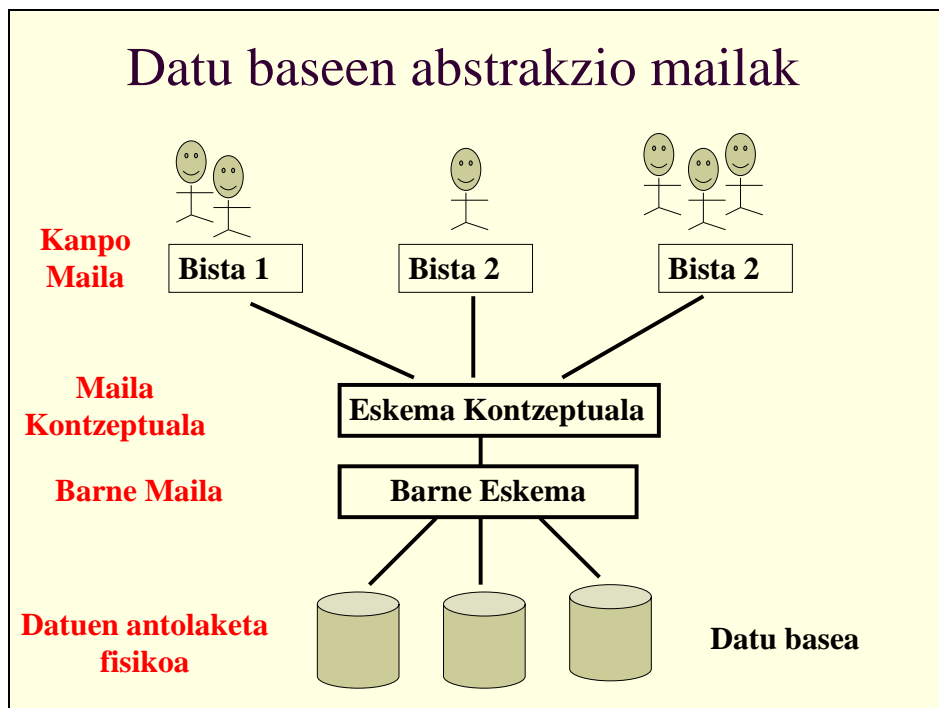


## 1. MODULUA: Informazioaren Antolaketa eta Egitura 1 Gaia: Datu baseen kontzeptua eta garrantzia.

### Datu base sistemen abstrakzio mailak

Datu base sistema baten lehenengo helburua erabiltzaileei informazioaren ikuspegi abstraktua ematea da. Horrela, erabiltzaileak ez du jakin behar datuak nola biltegituta dauden edo mantentze lana nola egin. Hori posible izateko, datu basea hiru abstrakzio mailatan banatzen da. Abstrakzio maila bakoitzean datuen deskripzioa egiteko, eskemak erabiltzen dira.

- **Maila fisikoa:** datuen biltegitatze fisikoaren deskripzioarekin zerikusia du. Maila honetan eskema fisikoa definitzen da.
- **Maila kontzeptuala:** datuen antolaketa logikoari lotuta dago. Maila honetan eskema kontzeptuala definitzen da.
- **Kanpoko mailak:** datuen inguruan erabiltzaileek dituzten ikuspegi deskripzioekin zerikusia du. Erabiltzaileengandik hurbilen dagoen maila da. Maila honetan kanpoko eskemak definitzen dira. Kanpoko eskema bat baino gehiago egon daiteke, bakoitza datu base osoaren zati baten adierazpen abstraktua izanik. Horrela, erabiltzaile bakoitzari, berearen beharrezan arabera dagokion informazioa erakutsiko zaio. Horrelako informazio multzo bakoitzari ikuspegi esaten zaio





## ***1. MODULUA: Informazioaren Antolaketa eta Egitura*** ***1 Gaia: Datu baseen kontzeptua eta garrantzia.***

### **Datu-baseak vs DBKS.**

Datu-basea deitzen zaio denbora-tarte batean existitzen den datu-bilduma egituratuari (adibidez, bideoklub bateko bazkideen datuak: izena, helbidea, telefonoa, etab.). Egungo datu-baseak ordenagailuei loturik daude eta haien kudeaketa guztiz automatizatuta dago, software berezien bidez. Software hori osatzen duten aplikazioek Datu-baseak Kudeatzeko Sistemak (DBKS) izenez ezagutzen da. Datu-baseak kudeatzeko sistema hori datu-base fisikoaren eta datu-base horren erabiltzaileen artean dagoen softwarea edo programen bilduma da. Erabiltzaileek datu-basea atzitzeko, bertako datuak irakurri, aldatu edo ezabatzeko erabiliko dute DBKS.

Teorian, datu-basea eta DBKS-a kontzeptu ezberdinak direnez, batetik software-etxe bateko datu-basea eta, bestetik, beste software-etxe bateko DBKS izan ditzakegu. Baina hori, normalean, ez da horrela izaten. Etxe bakoitzaren DBKS eta datu-basea ezin apur daitekeen bikotea da. Datuak formatu ezberdinetan gordetzen dituzte eta, ondorioz, eurek bakarrik dakite nola erabili. Erabiltzaileek estandarizatutako zenbait lengoia erabil ditzakete (SQL ezagunenetarikoa) datuekin lan egiteko eta lengoia horiek erabiliz egindako kontsultak DBKS bakoitzak nahi dituen bezala kudeatuko ditu, betiere gutxiengo ezaugarri batzuk betez. Horregatik, nahiz eta DBKS bakoitzak datuak modu ezberdinean erabili, erabiltzailea ez da ohartzen, berak DBKS-arekin soilik dituelako harremanak eta, horretarako lengoia estandarizatu batekin nahikoa duelako. Beraz, erabiltzailearen ikuspuntutik berdin dio zer DBKS/datu-base bikote erabili, komunikatzeko lengoia estandarizatu bat erabiliko baitu.

Azken finean, esan daiteke DBKS datu-basearekin lan egitea ahalbideratzen duen programa-bilduma dela.

DBKSak segurtasuna ere bermatu egin behar du; alde batetik, erabiltzaileen kudeaketaz eta atzipenentarako izango dituzten baimenez arduratuko da, eta bestetik, datuen segurtasunaz; datu basearen segurtasun kopiak egiteko tresnak izango ditu, istripuak gertatzen direnean datu basea egoera sendoan uzteko gai izan beharko du...

Datu baseek eta DBKsek bilakaera sendoa izan dute 60ko hamarkadan aurkeztu zirenetik gaur arte. Hasieran, eredu hierarkikoa zen nagusi; geroago, sare eredu agertu zen, eta azkenik, gaur egun merkatuan nagusi den eredu erlazionala. Egundak, ikerketak aurrera darraiela, datu base banatuak, Data Mining teknikak, objektuei zuzendutako datu baseak... egun badaude, baina oraindik merkatuaren gutxiengoa baino ez dira.





# 1. MODULUA: Informazioaren Antolaketa eta Egitura

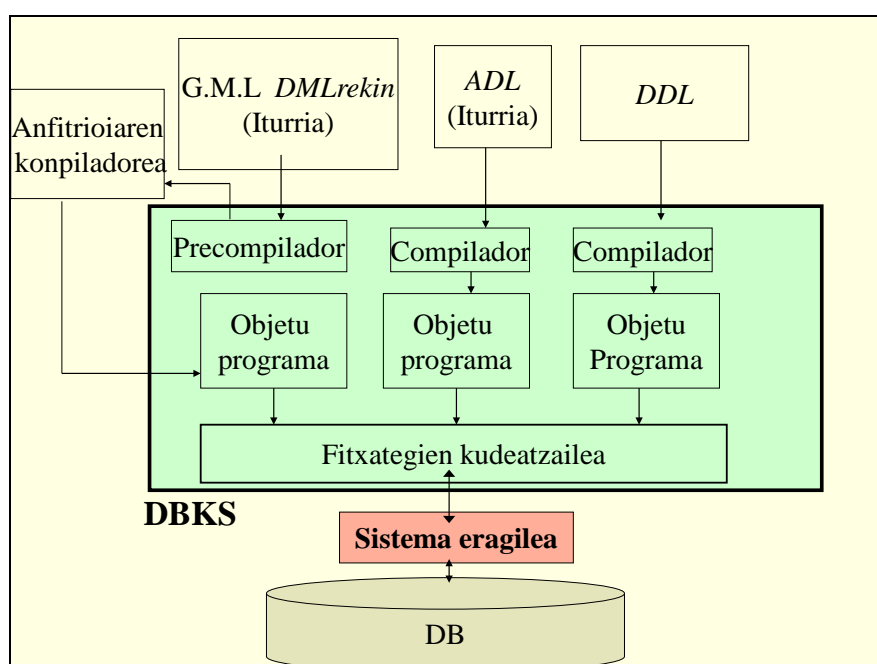
## 1 Gaia: Datu baseen kontzeptua eta garrantzia.

DBKSak, datuak definitu eta kudeatu ahal izateko, ondoren aurkezten diren bi lengoaiak erabiltzen ditu:

- Datuak Definitzeko Lengoaia (DDL): datu basea osatzen duten elementuak definitzeko erabiliko da. DDLa erabiliz egindako definizioak, murrizketak, osotasun arauak, datuen egiturak... datu hiztegia gordetzen dira.
- Datu Manipulatzeko Lengoaia (DML): datu basea definitu ostean, bertako datuekin lan egin ahal izateko, DBKSak ulertuko duen eragiketa multzoa behar da.

Eredu erlazionalean, DML eta DDL aginduak, denak egiten dira SQL lengoaia erabilita. Izan ere, datuen manipulazioa bi modutan egin daiteke:

- a) Modu zuzen batean. DBKS-ak datu basearekin “elakarrizketa zuzen” bat mantentzeko aukera ematen dio azken erabiltzaileari eskaintako interfaze edo lengoaiei esker (QBE, SQL, etab.).
- b) Modu diferidoan edo programazio bidez. Kasu honetan DML-a bi modutan erabili daiteke.
  1. Goi mailako programazio lengoaietan DML aginduak erantsiz.
  2. Datu baseetan oinarritutako aplikazioak garatzeko programazio lengoaia bereziak erabiliz (ADL ‘Application Development Language’)





***1. MODULUA: Informazioaren Antolaketa eta Egitura***  
***1 Gaia: Datu baseen kontzeptua eta garrantzia.***

Beste aldetik, DBKS-tan fitxategien gestioaz arduratzen den modulua fitxategien kudeatzailea da. Modulu hau sistema eragilearekin estu-estu komunikatzen da, maila baxueneko funtzioen egikaritzapenez arduratzen delarik, esaterako: erabiltzaile anitzen kontrolaz, datuen barne antolaketa fisikoaz, atzipen konkurrenteen gestioaz, segurtasun neurriak eskaintzeaz, etab.