



# Xenpelar Dokumentazio Jardunaldiak

## Bilaketa bakarra erabiltzaileentzat: datu guztiak klik batean

2019-05-10 Villabona, Subijana Etxea



### Sarrera

Xenpelar Dokumentazio Jardunaldi hauetan datu ezberdinak biltzen dituzten datu-baseak bilaketa bakarrean eskaintzen dituzten esperientziei buruz hitz egin da.

Ahozko ondareak berezko dituen ezaugarriengatik, hainbat gunetan biltzen diren datu eta dokumentazio mota ezberdinak modu bateratuan eskaintzeko beharra du, eta beharrizan hau ahozko ondaretik haragokoa ere bada. Interneten herritarrei ahalik eta informazio gehien eskuragarri jartzeko ondarearen digitalizazio eta datu-bilketa handiak egin dira azken urteotan, eta, ondorioz, gune ezberdinak sortu dira.

### Hitzaldiak

Iturri ezberdinetako datuak bilaketa bakarrean eskaintzeko esperientzia anitzak jaso dira jardunaldietan. Hitzaldi bakoitzean eta ondorengo mahai-inguruan aipatutako ideia nagusienak biltzen dira laburpen-dokumentu honetan:

- 📍 **Liburubila, online baliabideak eta mailegu sare zabala klik batean**  
Mikel Aizpuru. Eusko Jaurlaritzaren Ondare zuzendaria  
Francisca Pulgar. Eusko Jaurlaritzaren Liburutegi Zerbitzuko arduraduna
- 📍 **Artium museoko askotariko ondarea Interneten eskuragarri**  
Elena Roseras. Artiumeko Liburutegi arduraduna
- 📍 **Bilketa.eus, elkarlanaren emaitza euskal ondare bibliografikoan**  
Marie-Andrée Ouret. Baionako Mediateka. Euskal funtsen proiektuaren arduraduna
- 📍 **Solr bilaketa bateraturako motorea. Esperientziak**  
Aitor Maritxalar. Dokumentazioan espezializatutako informatika garatzailea

## **Liburubila, online baliabideak eta mailegu sare zabala klik batean**

Mikel Aizpuru. Eusko Jaurlaritzaren Ondare zuzendaria

Francisca Pulgar. Eusko Jaurlaritzaren Liburutegi Zerbitzuko arduraduna

**Hitzaldian erabilitako aurkezpena ikusi (pdf)**

*[Hitzaldian aipatutako ideia nagusien bilketa da hau]*

### **Egitasmoaren ezaugarriak.**

- Bi helburu nagusi:
  1. Ondarea herritarren esku jartzea, mota guztietako dokumentuak eskainiaz (testu, ikus-entzunezko; fisiko, digital; ...)
  2. EAEko liburutegien arteko sarearen identitatea sortzea, eta etorkizunean liburutegi gehiago sartzea. Urtero hazten doan sarea da.
- Oraindik prozesuan dagoen egitasmoa da, eta etorkizunean mundu digitala zabaltzeko asmoa dago.

### **Liburubilak bilatzaile bakarraren bidez eskaintzen diren lau datu-baseak:**

- [Katalogo Bateratua](#) (2004): [Absysnet](#) softwarearen bidez osatutako katalogo kolektiboa. EAEko udal liburutegiak eta sarean sartu diren bestelako liburutegien OPAC zerbitzua. *MARC21* formatua erabiltzen du.
- [Liburuklik](#) (2008): Irakurketa sare publikoko kide diren liburutegiek digitalizatutako liburu eta dokumentuak biltzen dituen gordailua, [Dspace](#) softwarearekin garatuta. *DUBLIN CORE* metadatuak erabiltzen ditu.
- [eLiburutegia](#) (2014): Sareko erabiltzaileentzat eskainitako liburu digitalen mailegu zerbitzua. *MARCXML* edo *DUBLIN CORE XML* formatuak darabiltza.
- [E-film](#) (2018): Sareko erabiltzaileentzat eskainitako film digitalen mailegu zerbitzua. Ez darabil liburutegietako formatu estandarrik.

### **Lau datu-base hauek liburutegietako beharrian ezberdinei erantzuna ematen die eta ezinbestean behar dira:**

- Liburutegien munduko teknologia garatzen hastean, katalogoak garatu ziren, liburutegietako lan prozesu guztiak egiten ditu (katalogazioa, mailegua, erosketak,...). Dokumentu fisikoen kudeaketarako tresna bezala garatu ziren.
- Gerora, dokumentuen digitalizazio edo numerizazio egitasmo handiak egin ziren, eta digitalizatutako objektuak (liburuak, aldizkariak, eta abar) kudeatzeko tresnak sortu ziren, gordailuak, objektu digitalen ezaugarri eta metadatuak modu estandarrean kudeatuz. [Europeana](#) gordailuan eskaintzen dira Europako liburutegi eta ondare zentroetan numerizatutako dokumentuak
- Ondoren, liburu digitalen mailegu zerbitzua sortu da. Argitaletxeek egileekin adostuta eskaintzen dituzte plataforma hauetan liburuak, eta adostutako epeak pasatutakoan aldaketak izaten dira. Mota hauetako plataformek aipatutako gai guztiak kudeatzen dituzte.
- Azkenik, filmak digitalki mailegatzeko zerbitzua garatu da. Kasu hauetan produktoreekin adosten dira filmak plataforma bidez eskaintzeko eskubideak eta aldaketa handiak izaten dira etengabe. Baliteke, egun batetik bestera 100 filma desagertu eta beste 200-500 inguru film berri agertzea e-film plataforman.

## Herritarrentzat ezinbestean bateratu behar. Metabilatzailearen hautaketa: [Vufind](#)

- Arrazoiak:
  - o Kode irekiko softwarea da.
  - o Belaunaldi berriko katalogoa deritzon softwarea da.
  - o Nazioarteko liburutegietan gehien zabaldutakoa da.
  - o Komunitate handia du atzetik, etorkizuneko garapena eta eguneratzea ziurtatuz.
- Ezaugarriak:
  - o Baliabide guztiak leku bakarrean kontsultatzeko aukera eskaintzen du.
  - o Interfazea hizkuntza ezberdinetan eskaintzen du.
  - o Fazeten bidezko bilaketak eta bilaketa iradokizunak eskaintzen ditu.
  - o Erabiltzaileak bere gogokoak eta zerrendak osatzeko aukera du.
  - o Erreferentzia bibliografikoen kudeatzaileekin bateragarria da, ikertzaileentzako erraztasunak eskainiaz.
  - o URL iraunkorrak ditu, internet bilatzaileen emaitzetan agertu ahal izateko.
  - o Beste informazio sistemekin elkar-eragin dezake, datuak eman eta jasotzeko.

### Datuen prozesua

- Vufind barruan hainbat zerbitzu daude, jasotako datuak bateratu, kudeatu eta erabiltzaileen esku interfaze bakarraren bidez jasotzeko. *[Ikusi hitzaldian erabilitako aurkezpena]*
- Zerbitzari bakarrean instalatuta dago, baina bi zerbitzari bezala antolatu da:
  - o Aurre-produkzioko zerbitzaria: Datuak nola gelditu diren ziurtatu eta zuzenketak eta datuen bateratzeak egin ahal izateko.
  - o Ekoizpeneko zerbitzaria: Erabiltzaileak jasoko duena, zuzenketa guztiak eginda.
- Datuen gehitzea lau iturrietatik:
  - o Egunero, Katalogo bateratuko (MARC21 formatuan) eta eLiburutegiko (XML formatuan) datu guztiak bidaltzen Vufind zerbitzarira eta datu hauen indexazioa prest uzten da goizean goiz.
  - o Liburukliken ez dago horrenbeste eguneraketarik. Behin egin zen datu guztien karga (DUBLIN CORE formatuan) eta orain objektu berriak gehitzen direnean bakarrik egiten da datuen karga Vufind-era.
  - o E-film plataformaren datuak ere kargatzen dira Vufind zerbitzarira, baina ez egunero, bakarrik aldaketa asko izaten baitituztenean. Datuak MRC fitxategi batean bidaltzen ditu Vufind zerbitzarira eta oraindik datuen bateratzea findu beharra dagoen arren, denak bilaketa bakarrean erakustea lortu da.
- Datu-garbiketa :
  - o Datuak bateratu egin behar dira, askotan erregistro edo egile berdina izan behar dena, bi eratan jaso delako (eta sistemak ez duelako identifikatu berdina dela).

### Hurrengo hobekuntza teknikoak:

- Erabiltzaileek Liburubila plataforman login egindakoa, online zerbitzu guztiak bertan egin ahal izatea lortu nahi da. Katalogo bateratuko Absysnet softwarearekin lortu da liburutegien arteko mailegu zerbitzua eta erabiltzaileen datuak (itzultzeko datak, erretserbak eta abar) Liburubila bidez eskaintzea. Etorkizunean eLiburutegia eta e-film mailegu digitaleko zerbitzuak ere modu bateratuan eskaintzea da helburua.

## **Artium museoko askotariko ondarea Interneten eskuragarri**

Elena Roseras. Artiumeko Liburutegi arduraduna

### **Hitzaldiaren laburpen osoa irakurri hemen (pdf)**

*[Hitzaldian aipatutako ideia nagusien bilketa da hau]*

#### **Egitasmoaren ezaugarriak.**

- Helburu nagusia: Kultur ondare garaikidearen ezagutza eta erabilera artearen, hezkuntzaren, ikerketaren munduan, eta orokorrean herritarren artean bultzatzea.
- Helburuak:
  1. Museoak sortutako dokumentazio eta informazio guztia modu eraginkorrean kudeatzea
  2. Artiumen ondarea zerbitzu publiko bakar baten bidez eskainiz, museoak sortutako jakintza guztia hobeto ezagutarazi eta zabaltzea.
- Museoko sail ezberdinetan sortutako informazioa eta dokumentazioaren arteko harremana egituratu da.

#### **Artiker egitasmoa abiatu da, bilatzaile bakarraren bidez datu-base guztiak elkarren artean lotzeko**

*[Xehetasun gehiago hitzaldiaren laburpen dokumentuan]:*

- **Liburutegia edo funts bibliografikoak** (museoak egindako ikus-entzunezko grabazioak, prentsa artikuluen datu-basea, eta artxibo espedienteak ere barne). Dokumentu fisikoez gainera, digitalizatutako dokumentuak ere badira. Katalogaziorako, **Amicus-LibriSuite** softwarea; eta erabiltzaileei online eskaintzeko, **Vufind** softwarea. Fitxategi digitalak Scanbitek garatutako bisore batekin egokitu dira.
- **Museoaren jardueren datu-basea**. **Access** softwarearekin garatuta. Erakusketak, hitzaldiak, ikastaroak eta lantegiak, zinema-zikloak, kontzertuak, eszena garaikidea eta abar. Jarduera guztien fitxetan jaso dira beste jarduerekin edo museoko beste atalekin (webguneko edukiak, artista edo erakusketen informazioa, liburutegiko dokumentuak,...). Gaur egun datu-base honetako edukia Amicus-Librisuiten bidez katalogatu eta interneten Vufind bidez erakusten da.
- **Argazki artxiboa**. Portfolio softwarearekin kudeatuta. Argazki guztiak digitalak dira (jpg edo tiff), baina batzuen jatorrizko formatua papera da, edo negatiboak dira, eta digitalizatu egin dira. Gaur egun argazkiak Amicus-Librisuiten bidez katalogatu eta interneten Vufind bidez erakusten da; aurreko bilduma berrantolatu, garbitu eta modu honetara egokitu da.
- **Ondare artistikoa**. **EMSIME** sistemaren bidez kudeatuta. EMSIME Eusko Jaurlaritzak museoetako ondarea kudeatzeko garatutako katalogo kolektiboa da. EMSIME plataformatik eremu garrantzitsuenak excel-en esportatu eta XML LIDO sistemara moldatzen dira, informatikariek egiten dute azken moldaketa hori. Lehen kontsulta baterako garrantzitsuenak izan daitezkeen eremuak bakarrik esportatu dira, eta erabiltzaileak informazio zabalagoa lortu nahi badu, EMSIMEra daraman esteka baten bidez egin dezake.
- **DokuArt, erakusketen artxiboa**. **Drupal** softwarearekin garatuta. Museoak sortutako txostenak kudeatzen ditu, eta jardueren inguruan sortutako informazio guztia eskaintzen da: azalpen testuak, biografiak, eta abar.

#### **Elkarren artean lotu eta modu bateratuan erabiltzaileei eskaintzeko erabakia:**

- **Finna** egitasmoa hartu da eredu: Finlandiako liburutegi, museo eta artxiboetako datu-base eta katalogoak modu bateratuan eskaintzen duen bilatzailea. **Vufind** softwarea erabiltzen du.
- Bi fasetan egin zen:
  1. Lehenengo, katalogo bibliografikoaren bertsioa eguneratu (Amicus-LibriSuite), jardueragordailuaren eta argazki-artxiboaren egitura finkatu zen
  2. Bigarren fasean, ondare artistikoa erantsi zen proiektuari.

## [Bilketa.eus](http://Bilketa.eus), elkarlanaren emaitza euskal ondare bibliografikoan

Marie-Andrée Ouret. Baionako Mediateka. Euskal funtsen proiektuaren arduraduna

### [Hitzaldian erabilitako aurkezpena ikusi \(pdf\)](#)

*[Hitzaldian aipatutako ideia nagusien bilketa da hau]*

#### Egitasmoaren ezaugarriak.

##### - Elkarlana du oinarrian:

1. Egitasmoa bultzatzen duten erakundeek elkarlana lortu da:
  1. Baionako Herria – Baionako Mediateka
  2. Euskal Hirigune Elkargoa
  3. Pirinio Atlantikoetako Departamendua
  4. Akitania Berriko Eskualdea
  5. Frantziako Kultur Ministerioa (Akitaniako DRAC erakundearen bitartez)
  6. Frantziako Biblioteka Nazionala
2. [Funtsak eskaini dituzten erakundeek arteko elkarlana](#) lortu da, sare bat osatuz:
  1. Herri liburutegiak
  2. Unibertsitateko liburutegiak
  3. Artxibategiko liburutegia
  4. Ikerketa zentroa
  5. IKAS Pedagogia Dokumentazio zentroa
  6. Euskal Museoa
  7. Apezpikategiko liburutegia

##### - Helburua: Ipar Euskal Herriko euskal funtsak modu bateratuan eskaini herritarrei.

1. Euskaraz idatzitako guztiak
2. Euskal gaiak lantzen dituztenak: euskara, Euskal Herria, euskaldunak,...

##### - Informazio mota guztiak batera:

1. Edozein formatu: eskuizkribu, liburu, aldizkari, artikulua, ikus-entzunezko, mapa, ...
2. Fisikoak eta numerikoak batera:
  - a. Dokumentu fisikoaren deskribapen fitxak (130.000 erreferentzia)
  - b. Dokumentu integralak: objektu numerikoak/digitalak eta beraien deskribapen fitxak (1.500 dokumentu numeriko)

#### Informazio iturriak:

- Informazioa eskaintzen duten erakundeak: Euskal funtsen eskuragarritasuna lortzeko helburuarekin Izaera ezberdineko erakundeekin sare bat osatu eta elkarren arteko elkarlana bultzatu da. Euskal funtsen katalogaziorako tresna bateratuak sortzeko elkarlana ere antolatu da.
- Numerizazio egitasmoa: Frantziako Liburutegi Nazionalaren eta Baionako Mediatekaren arteko elkarlanari esker jarri da abian, eta Bilketako iturri diren erakundeetan dauden euskal funtsak numerizatzeko egitasmoa da. Numerizatu behar diren liburu eta prentsa bildumak daude hainbat erakundeetan.

#### Frantziako Biblioteka Nazionalekiko elkarlana - Etorkizun ikuspegia:

[Frantziako Biblioteka Nazionalean \(BnF\)](#) dauden euskal funtsak etorkizunean jaso:

- Ugariak eta oso garrantzitsuak dira (zaharrenak, bereziak,... 8.000 dokumentu inguru)
- Numerizatutako funts handia da: 700 dokumentu [Gallican](#).
- BnF-rekin elkarlanean, euskal funtsa bereizteko kodeketa bat lortu da, datuak Bilketara ekarri ahal izateko.

## Datuen prozesua

Sistemak dituen antolaketa [Ikusi hitzaldian erabilitako aurkezpena]:

- Gordailuan jasotzen dira datu guztiak eta OAI uzta bidez Bilketan eskaini:
  - o Kanpo iturrietatik jasotako datuak.
  - o Liburutegi numerikoa, objektu digitalekin.
  - o Autoritateak.
- Formatuak:
  - o Inportatutako formatuak: UNIMARC (datu bibliografikoak) eta EAD (ISAD-G) edo XML (artxibo edo eskuizkribuen datuak).
  - o Autoritateak: Formatu propioan eta UNIMARC esportaziorako
  - o Eskaintzen dituen formatuak: DUBLIN CORE (DC) eta DUBLIN CORE-FB
- Bilketan Gordailuko OAI uztarekin batera, beste OAI uztak ere gehitzen dira.
- Metabilatzailea: [SOLR](#) motorea. Bilaketa fazeten bidez erabiltzaileari laguntzeko.

## Datuen bateratzea

Iturri ezberdinetatik jasotako datuen bateratze eta garbiketa handia egiten da:

- Data fazeta: urteak eta mendeak eskaintzeko
- Autoritateen datu-basea, onartutako formak eta onartu gabeak finkatzeko. Euskal autoritateen formak finkatzeko lan-taldea eratu da.
  - o Egileen forma ezberdinak
  - o Leku izenak (ez da oraino landua)
- Iturrietako inportazioak etengabe zaindu eta konpondu beharra: Aldaketa txiki bat gertatzen denean iturri bateko datutegian, eragin zuzena du OAI uztan.

## Bilketa egitasmoaren fase berria

Sistemaren teknologia berrituko da hurrengo urteetan: [Bibliomondo](#) enpresak egingo du Bilketaren teknologia berritzea.

## Lortu nahi diren hobekuntza teknikoak:

- OAI uzta: metadatuak sakondu:
  - o BnF-ko APIak jasotzeko (erregistroak, dokumentu numerikoen metadatuak eta edukiak – testu, alto fitxategiak, irudiak,...).
  - o Datuak zuzenean tratatzeko/zuzentzeko moduak
  - o Dublin Core metadatuaren bilketa optimizatu
- Bilaketa motorea indartu:
  - o Indexazioa aberastu:
    - testua, barne eta kanpoko erakundearen dokumentu numerikoetan
    - Bilketa webguneko edukietan
    - Autoritateetan, ez onartutako formak
  - o Emaitzen bistaratzea landu:
    - Faketak, sailkapenak, fulltext emaitzak
    - Prentsa numerikoaren bisualizazioa
    - Artxibo dokumentuen hierarkiaren bisualizazioa
  - o Informazio ugartasuna kudeatu
- Autoritateen kudeaketa indartu (Egileak lehen, gai hitz gakoak edo deskriptoreak etorkizunean,..):
  - o Identifikazio automatikoak egiteko aukerak indartu
  - o Onartu ez diren formen kudeaketa eta bilaketa indartu
- Eduki numerikoak: nabarmendu eta HD irudiak sisteman integratu

## **Solr bilaketa bateraturako motorea. Esperientziak**

Aitor Maritxalar. Dokumentazioan espezializatutako informatika garatzailea

### **[Hitzaldian erabilitako aurkezpena ikusi \(pdf\)](#)**

*[Hitzaldian aipatutako ideia nagusien bilketa da hau]*

### **Solr bilatzailearen ezaugarriak** [*Xehetasun gehiago aurkezpen dokumentuan*]

- Testuan oinarritzen da eta dokumentua oinarri: ez da erlazionala.
  - o Testu identikoak bateratuko ditu, erregistroen arteko harremanak eratuz.
  - o Testua zentzu zabalean: datu-baseetako esportazioko datuak; pdf edo bestelako dokumentuetako testuak; argazkien metadatuak; ...
- Bi motore: informazioa indexatzeko motorea eta informazioa berreskuratzeko motorea
- Kode irekian banatzen da (Apache lizentzia). Doan erabili eta moldatu daiteke
- Datu-base eskema malgua du: beste fulltext bilatzaileetan testua bakarrik sartu daiteke, eta hemen, testu egituratua – dokumentua deritzona - ere sartu daiteke (garrantzitsua dokumentazioaren munduan, testu planoak kudeatu ordez, norberak finkatutako eremuetan antolatu ahal izateko, eta eremu bakoitzari dagokion garrantzia eman).
- Oso azkarra eta egonkorra da, hizkuntza ezberdinetako bilaketa ezaugarriak kudeatu ditzake, eta pluginak ditu behar diren ezaugarrietara egokitzeko.
- Erabiltzailearentzat, bilaketak egiteko eta emaitzak bistaratzeko erraztasun handiak eskaintzen ditu

### **Bilaketa motoreak eta softwareak**

- Solr bilaketa motoreak software bat beharko du formularioak eta webgunea garatu eta datuak erakusteko. Webaren garapena apartekoa da (aldatzeko ere motorean eraginik ez). [*Ikusi indexazio prozesuaren eskema aurkezpenean*]
- Gure beharren arabera aukeratu behar dira motorea eta softwarea.
- Bilaketa motorea + softwarea biak integratuta eskaintzen dituzten aukerak daude:
  - o [VUFIND](#): Datu bibliografikoen kudeaketarako garatutako metabilatzailea eta web nabigatzailea. Solr dauka integratuta.
  - o [SAMVERA](#): Gordailu moduan garatuta, artxiboetako edukiak kudeatzeko egokia izan daiteke. Solr dauka integratuta.

### **Hautaketa gure beharrak ondo aztertuta egin behar da**

- Lehenengo gure dokumentazioaren ezaugarriak eta lortu nahi diren helburuak ondo aztertu. Ondoren aukeratu komeni den motorea edo/eta softwarea.
- Solr erabiltzea noiz aukeratu (beste software batean integratu gabe):
  - o Eduki oso heterogeneoak batu nahi direnean, ez datu bibliografikoak bakarrik; edo eduki oso berezia denean (fazeta bereziak)
  - o Indexazio eta bilaketa emaitzen kontrol osoa nahi denean

**Motore eta/edo softwarearen hautaketa prozesua. 4 urratsak** (bi esperientzien ezaugarriak alderatu dira: [BDB](#) eta [euskaltzaindia.eus](#). *Ikusi aurkezpen dokumentua*):

- Beti kontutan izan: bilatzailea diseinatzean, erabiltzailea buruan:
  - o Ez du ezagutzen gure datu eta informazioaren egitura; ezta gure antolaketa.
  - o Normalean ez daki zehazki zer bilatzen duen.
- 1. Bilatzailearen helburua:
  - Bilaketa egiteko eremu bakarra ala bilaketa gidatua? Edukien ikusgarritasuna lortu nahi da? (salmenta) Eduki konkrituak ala material gehien erakutsi nahi da?
- 2. Erabiltzaile profila definitu:
  - Estatistikak begiratu eta zein ataletan dago erabilera gehien ziurtatu. Erabiltzaile espezializatua edo orokorragoa den identifikatu.
- 3. Gure bilatzailearen prezisioa eta estaldura finkatu
  - Biak %100ean izatea ezinezkoa da: bata igotakoan, bestea jaisten da:
    - Prezisioa: bilatutakoa bakarrik ematea
    - Estaldura: bilatutako guztia ematea
  - Orokorrean:
    - Erabiltzaile orokorra bada: prezisioa jaitsi eta estaldura igo: emaitza gehiago etorriko dira. Ez du esan nahi emaitza txarrak direnik. "Positibo faltsuak" izatea onargarria da, emaitza guztiak barruan daudela ziurtatuz eta deskubrimenduzko edukiak ere eskainita.
    - Erabiltzaile espezialista: prezisioa igo eta estaldura jaitsi → Bilatzen ari dena gehiago daki, eta deskubrimenduko emaitzak ez da horrenbeste behar.
- 4. Indexatu nahi diren baliabideen inbentario zehatza egin:
  - Zeintzuk datu-base, bildumatan dauden bilatu behar diren datu guztiak
  - Zeintzuk formatu ditugun: datuak, pdf, word, irudien metadatuak,...
  - Zeintzuk hizkuntza dauden jasota
  - Kopuruak: erregistroak, dokumentuak,...

#### **Solr motorea ezartzeko egin beharreko lanak:**

1. Dokumentua (schema) diseinatu informatikariarekin batera:
  - Bilatzailea diseinatzerako orduan urrats garrantzitsuenetakoa da. Jasoko diren datuak zeintzuk eremu eta eremu motetan antolatuko diren erabaki. Mota bakoitzaren garrantzia ere erabaki.
  - Facet-ak: oso garrantzitsuak. Emaitzak iragazteko izateaz gain, bilaketa-deskubrimendua egiteko aukera ematen dute, gure informazioaren egitura azaltzen dute, eta dimentsioa, kopuruak eskaintzen dute.
2. Esportazioak eta garbiketak egin: JSON fitxategi erraldoiak sortzen dira bildu nahi diren baliabide guztietako datuekin eta aurrez definitutako schema horretan jasoko dira. Jasotako datuen bateratzeak eta zuzenketak egin behar dira, bilaketa emaitzak eta fazetak behar diren moduan egokitzeko.
3. Stemmera edo [lematizazioa](#): Euskararen kasuan garrantzi berezia du, baina ondo frogatu behar motorrak jasotako datuekin. Estaldura asko igotzen du, baina prezisioa jaitsi. Testu zaharrekin ez da komeni.
4. Solr bilatzailea dagokion webgunearekin lotu.



## Antolatzaileak

2019ko Xenpelar Dokumentazio Jardunaldien antolatzaileak Mintzola Ahozko Lantegia, Bertsozale Elkartea eta Iametzta Interaktiboa gara. Artxibategi, Liburutegi eta Dokumentazio Zentroetako Profesionalen Euskal Elkarrean (ALDEE) lagundu du jardunaldien berri zabaltzen.

Antolatzaileak:



Laguntzaileak:

