

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Karrera Bukaerako Proiektua

2013ko otsailak 6

Iñigo Lopez Gazpio

Zuzendaria:

Montse Maritxalar Anglada

eman ta zabal zazu



Universidad del País Vasco Euskal Herriko Unibertsitatea

Gaien Aurkibidea

1	Sarrera	1
1.1	Sarrera	2
1.1.1	Aplikazioaren izenaren zergatia	2
2	Proiektuaren helburu dokumentatua	3
2.1	Proiektuaren deskribapena eta helburuak	4
2.2	Proiektuaren abiapuntua	6
2.2.1	Proiektura egokitzeko egindako aldaketak	7
2.3	Ataza eta azpiatazen zerrenda	8
2.4	Lanaren Deskonposaketa Egitura (LDE diagrama)	9
2.5	Proiektuaren plangintza	11
2.5.1	Estimazioa	11
2.5.2	Gantt diagrama	12
2.6	Lan metodologia	17
2.7	Bideragarritasuna	18
2.8	Arriskuen Analisia	18
2.8.1	Arriskuen zerrenda	19
2.8.2	Arriskuen kontingentzia plana	19
3	Analisia	21
3.1	Erabiltzaileen beharren eta antzeko tresnen azterketa	22
3.1.1	Ondorioak	24
3.2	Aplikazioaren arkitektura	26
3.3	Erabilpen kasuen analisia	30
3.3.1	Erabiltzaile kautotu gabeen erabilpen kasuen analisia	31
3.3.2	Erabiltzaile kautotuen erabilpen kasuen analisia	32
3.3.3	Administratzailearen erabilpen kasuen analisia	33
3.4	Software erabakien eta teknologiaren analisia	34
3.4.1	Erabilitako azpiegitura, teknologia eta baliabideak	34
3.4.2	IXA taldean garatutako tresnen eta baliabideen erabilera	40
3.4.3	Frameworken inguruko erabakia	41

3.4.4	Txostena LaTeX bidez garatzeko erabakia	42
4	Diseinua	43
4.1	Galdera sortzailea	45
4.2	Lehen bertsiotik bigarren bertsiora	50
4.2.1	Sarrera corpusaren tratamendua egokitzen	50
4.2.2	Galderak eta zehar galderak baztertzen	53
4.2.3	Hautagaien aukeraketa egokitzen	53
4.2.4	Galdetzaileen esleipena egokitzen	56
4.2.5	Galderaren eraikuntza egokitzen	59
4.3	Web aplikazioaren diseinua	61
4.3.1	Deskribapena	61
4.3.2	Modulu nagusien diseinua	63
4.4	Galdera sortzaile eta web aplikazioaren arteko komunikazioa .	75
4.4.1	Sarrera corpusa analizatzeko prozesua	75
4.4.2	Galderak sortzeko prozesua	76
4.5	Informazioa egituratzea eta datu basearen diseinua	80
4.6	Erabilpen kasuen diseinua	84
4.6.1	Erabiltzaile kautotu gabeen erabilpen kasuen diagrama	84
4.6.2	Erabiltzaile kautotuen erabilpen kasuen diagrama . . .	86
4.6.3	Administratzailearen erabilpen kasuen diagrama	89
4.6.4	Erabilpen kasuen fluxu diagramak	93
4.7	Segurtasuna bermatzen	102
4.7.1	Fitxategi sistema babesten	103
4.7.2	Pasahitzak babesten	104
4.7.3	Erabiltzaileen fitxategi pribatuak babesten	105
4.7.4	SQL injection-en aurka babesten	108
5	Software garapena	111
5.1	Softwarea garatzeko prozesua	112
5.1.1	Garapen iteratibo inkremental prototipatua	112
5.2	Web aplikazioa munduratzea	113
5.2.1	Search Engine Optimization	114
6	Ebaluazioa, emaitzak eta arazoak	117
6.1	Galdera sortzaile sistemaren ebaluazioa	118
6.1.1	Emaitzak	118
6.1.2	Errore analisisa	120
6.1.3	Ondorioak	120
6.2	Web interfazearen ebaluazioa	120
6.2.1	Evaluaketaren nondik norakoak	122

6.2.2	Erabiltzaileen iritzia	122
6.2.3	Ondorioak	125
6.3	Erabiltzaileen feedback-a	125
6.4	Proiektuan zehar izandako arazoak	127
7	Proiektuaren birplanifikazioa	131
7.1	Birplanifikazioa	132
7.2	Atzerapenak eta denbora desbideraketak	135
8	Ondorioak eta etorkizuneko lana	137
8.1	Ondorioak	138
8.1.1	Proiektuaren ondorioak	139
8.1.2	Ondorio pertsonalak	140
8.2	Etorkizuneko lana	141
8.2.1	Epe motzean egitekoak	141
8.2.2	Epe luzean egitekoak	142
9	Bibliografia	143
9.1	Liburuak	145
9.2	Artikuluak	145
9.3	Interneteko erreferentziak	147
I	Eranskinak	149
I.a	Analisi linguistikoan arazoak	150
I.b	SQL injection-en adibide erraz bat	153
I.c	Galdera sortzaile sistemaren ebaluaketako emaitzak	154
I.d	Eskuliburua	159

Irudien zerrenda

2.1	Lanaren Deskonposaketa Egitura (LDE diagrama).	10
2.2	Proiektuaren denbora plangintza.	12
2.3	Gantt diagrama.	14
2.4	Gantt diagrama.	15
2.5	Gantt diagrama.	16
3.1	ikasten.net logotipoa.	23
3.2	Bezero zerbitzari ereduaren eskema.	26
3.3	Geruzetan banatutako programazio ereduaren eskema.	28
3.4	MVC ereduaren eskema.	29
3.5	Aplikazioaren arkitekturaren eskema.	30
3.6	Ajax exekuzio eskema eredu.	37
4.1	Eraldaketa erregelaren errepresentazio grafikoa.	48
4.2	Galdera sortzailearen hasierako eskema logikoa.	49
4.3	Web aplikazioaren diseinuaren eskema.	66
4.4	Sarrera orrialdearen interfazea.	67
4.5	Sarrera orrialdearen interfazea erabiltzaile kautotuentzat.	68
4.6	Erabiltzaileak kudeatzeko interfazea.	69
4.7	Erabiltzaileak kudeatzeko interfazea.	70
4.8	Ariketak lantzeko interfazea.	71
4.9	Fitxategiak kudeatzeko interfazea.	72
4.10	Galderak zuzentzeko -editatzeko- interfazea.	73
4.11	Galderak erantzuteko -lantzeko- interfazea.	74
4.12	Web aplikazioak corpusak analizatzeko eta galderak sortzeko jarraitzen duen fluxu diagrama.	78
4.13	Web aplikazio, erabiltzaile, analizatzaile sintaktiko eta galdera sortzaile sistemaren arteko komunikazio eskema.	79
4.14	datu baseko entitate erlazio diagrama.	83
4.15	Erabiltzaile kautotu gabeen erabilpen kasuen diagrama.	85
4.16	Erabiltzaile kautotuen erabilpen kasuen diagrama.	87

4.17	Erabiltzaile kautotuen erabilpen kasuen diagrama.	88
4.18	Administratzailearen erabilpen kasuen diagrama.	90
4.19	Administratzailearen erabilpen kasuen diagrama.	91
4.20	Administratzailearen erabilpen kasuen diagrama.	92
4.21	Pasahitza berreskuratu erabilpen kasuaren fluxu diagrama. . .	93
4.22	Profila ikuskatu edo aldatu erabilpen kasuaren fluxu diagrama.	94
4.23	Kontua ezabatu erabilpen kasuaren fluxu diagrama.	95
4.24	Txatean parte hartu erabilpen kasuaren fluxu diagrama. . . .	96
4.25	Testu fitxategia igo erabilpen kasuaren fluxu diagrama.	97
4.26	Galderak zuzendu erabilpen kasuaren fluxu diagrama.	98
4.27	Zuzenketak gorde erabilpen kasuaren fluxu diagrama.	99
4.28	Ariketak landu erabilpen kasuaren fluxu diagrama.	99
4.29	Automatikoki zuzendu erabilpen kasuaren fluxu diagrama. . .	101
4.30	Erabiltzaileen fitxategiak ikusi erabilpen kasuaren fluxu dia- grama.	102
4.31	Pasahitzak biltegitratzeko eskema.	105
4.32	Kautotzea burutzeko eskema.	105
4.33	Erabiltzaileen fitxategiak babesteko ebazpenaren diagrama. . .	107
4.34	SQL injection ekiditeko eskema.	109
5.1	Softwarea garatzeko iterazioen eskema.	113
5.2	Google Analytics web kudeatzailea.	115
7.1	Birplanifikazioaren Gantt diagrama.	133
7.2	Birplanifikazioaren Gantt diagrama.	134
I.1	Ariketak lantzeko txokoan klik egin.	159
I.2	Webgunearen sarrera orrialdea.	160
I.3	Fitxategia aukeratu eta irakurri botoia sakatu.	161
I.4	Ariketak landu botoia sakatu.	161
I.5	Ariketak lantzeko interfazea.	162
I.6	Galderak erantzuteko ingurunea.	163
I.7	Erabiltzaile kontua sortu.	164
I.8	Erabiltzaile kontua sortu.	165
I.9	Funtzionalitate berri bat agertzen da kautotzean.	166
I.10	Fitxategiak kudeatzeko txokoa.	167
I.11	Galderak zuzendu botoia sakatu.	168
I.12	Galderak editatzeko interfazea.	169

Taulen Zerrenda

2.1	Proiektuaren denbora plangintza.	11
3.1	Erabiltzaile kautotu gabeen beharrak identifikatzen dituen taula.	31
3.2	Erabiltzaile kautotuen beharrak identifikatzen dituen taula. . .	32
3.3	Administratzailearen beharrak identifikatzen dituen taula. . .	33
4.1	Aditz jokatu eta jokatu gabeak identifikatzeko eskema.	46
4.2	Hautagaien aukeraketa heuristikoa. Lehen bertsioan	48
4.3	Alegiazko elementuak sortzeko eskema.	53
4.4	Hautagaien aukeraketa heuristikoa. Bigarren bertsioan	56
4.5	Galdetzaileen zerrenda.	57
4.6	Erabiltzaile, partekatze eta atzipen mota ezberdinak erlazio- natzen dituen taula.	106
7.1	Proiektuaren denbora estimazioak vs errealitatea.	136

Laburpena

Seneko testuak lantzeko web bidez eskaintzen den aplikazio didaktikoa da. Oro har, sistemak erabiltzaileengandik jasotako fitxategien gainean automatikoki galderak sortzen ditu eta ariketa gisa eskaini. Honetaz gain, ikaskuntza/irakaskuntza metodologiari dagokionez sarean aurki daitezkeen aplikazioen alternatiba gisa garatu da sistema. Izan ere, fitxategiak eta ariketak partekatzeko metodoak oinarritzat hartuz kooperazioa eta elkarlana ahalbidetzen duen metodologiari heltzea izan da helburua. Hau guztia gauzatzeko, besteak beste, web aplikazioak galderak automatikoki sortzeko gai den sistema batekin elkar eragiten du.

Halaber, sistema ebaluatzeko burututako saioetan jasotako informazio guztia biltzen da; bai galdera sortzailea ebaluatzeko hizkuntzalariekin egindako esperimentukoa, baita Martuteneko Arantzazuko Ama eskolan egindako laborategi saiokoa ere.

1 Kapituluia

Sarrera

Gaien Aurkibidea

1.1	Sarrera	2
1.1.1	Aplikazioaren izenaren zergatia	2

1.1 Sarrera

Ariketak automatikoki sortzen dituzten aplikazioak ikaskuntza/irakaskuntza arloekin erabat loturik dauden baliabideak dira.

Gaur egun oso ohikoa da arlo espezifikoak lantzeko erabiltzen diren aplikazioak sarean atzigarri aurkitzea. Era honetan, irakaskuntza, autoikaskuntza edota ikaskuntza ez presentziala ahalbidetzen dutelarik, eta, gainera, abantaila asko eskaintzen dizkiote bai erabiltzaileari baita irakasleari ere. Esaterako, erabiltzaileak ikasketari eskaintzen dion denbora bere beharretara egokitu dezake, eta ez hori bakarrik, ikasketa presentzialak dituen arazo geografikoak ere guztiz ekidin ditzake.

Ikaskuntza/irakaskuntza software inguruneek beste helburu garrantzitsu bat ere badute: irakaslearen lan karga txikitzea. Gainera, ariketak automatikoki eskaintzeko gai direnez, materiala prestatzeko eta osatzeko beharrezkoa den denbora tartea laburtzen dute. Buruturiko ariketak zuzentzeko, aldiz, aukera bi izaten dira erabilienak: automatikoki egikaritzea edota ikastaroaren arduradunak egitea eta zuzendu ostean, ikasleari emaitzak bueltatzea.

Honetaz gain, beste abantaila aipagarri bat ere atxikitzen zaie sarean oinarritutako ikaskuntza aplikazioei, izan ere, maila geografiko zabalagoan kooperatzeak kalitate eta aniztasun handiagoko materiala lortzea ahalbidetzen du.

Proiektu honetan aurkezten den lana euskara ikasi nahi duen edonori baliabide berri bat eskaintzeko sortu da. Horregatik, aipatu beharra dago ez duela inondik ere euskaltegirik ordezkatu nahi, ez eta euskara lantzeko erabiltzen den beste tresnarik. Hala ere, kooperazioa eta elkarlana bultzatuz ikuspuntu desberdin bat eskaintzen duen tresna izatea espero da handik kanpo euskara landu nahi duen edonorentzat.

1.1.1 Aplikazioaren izenaren zergatia

Garatu den aplikazioari Seneko izena eman izanaren arrazoi nagusiak hiru dira. Batetik, garatutako galdera sortzailearen helburu nagusietako bat ahalik eta galdera zentzudunenak egitea izan delako. Bestetik, jakitunak zentzuz jokatzen dutelako eta aplikazio honen helburua erabiltzaileak trebatzea delako. Azkeneko argudioa izenarekin erlazionaturik dago, Seneko hitz gozo eta leuna izanik, erabiltzailearengan onurazko eragina sortuko duela uste baita.

2 Kapituluia

Proiektuaren helburu dokumentatua

Gaien Aurkibidea

2.1	Proiektuaren deskribapena eta helburuak	4
2.2	Proiektuaren abiapuntua	6
2.2.1	Proiektura egokitzeko egindako aldaketak	7
2.3	Ataza eta azpiatazen zerrenda	8
2.4	Lanaren Deskonposaketa Egitura (LDE diagrama)	9
2.5	Proiektuaren plangintza	11
2.5.1	Estimazioa	11
2.5.2	Gantt diagrama	12
2.6	Lan metodologia	17
2.7	Bideragarritasuna	18
2.8	Arriskuen Analisia	18
2.8.1	Arriskuen zerrenda	19
2.8.2	Arriskuen kontingentzia plana	19

Proiektuaren helburu dokumentatua garatu den aplikazioaren azalpen zehatza egingo da. Proiektuaren deskribapena eta helburuak azaltzeaz gain, plangintza estimazioa, ataza eta azpiatazen zehaztapena, bideragarritasuna, arriskuen analisi eta kontingentzia plana eta lan metodologia azalduko da.

2.1 Proiektuaren deskribapena eta helburuak

Seneko sarean eta interesa duen edozein erabiltzailerentzat eskaintzen den e-learning aplikazio bat da. Oro har, e-learning aplikazioak urruneko ikaskuntza/irakaskuntzan aritzeko erabiltzen dira. Garatu den web aplikazio honen bitartez euskaraz idatzitako testuak lantzea ahalbidetzen da. Honetarako, galdera sortzaile (QG) sistemak testu baten gaineko galderak sortuko ditu automatikoki eta web aplikaziora esportatuko dira.

Hortaz bada, aurkezten den proiektuaren helburu nagusia bikoitza da: alde batetik, QG sistema garatzea, eta bestetik, web aplikazioa garatzea.

Alde batetik, galdera sortzaile sistemari dagokionez, garapena ez da ezezetik hasiko, aurretik ikasleak lan poltsa batean garatutako lana baitu abiapuntu. Proiektu honetan, lan poltsan garatutako QG sistemaren lehen bertsioa hartu, eta hurbilpen linguistiko berriak erabiliz, lehen bertsio honen gainean lan egingo da bigarren bertsio bat lortzeko. Honetarako, galderen sorkuntzan parte hartzen duten faseak analizatu eta egokituko dira. Besteak beste, sarrera corpora prozesatzea, analizatzaile sintaktikoaren akatsak leuntzea, galdera eta zehar galderak baztertzea, hautagaien aukeraketa optimizatzea, galdetzaileen esleipena egokitzea eta galderaren eraikuntza gaiak jorratuko dira; honetarako, bai Natural Language Processing (NLP) baliabideak aztertzea, baita XML markatze lengoaien zehaztapenak aztertzea beharrezkoa izango da.

Beste alde batetik, hutsetik garatu den web aplikazioari dagokionez, gaur egun sarean aurki daitezkeen ikaskuntza/irakaskuntza aplikazioen alternatiba gisa aurkeztu nahi izan da, webgune gehienek eduki estatiko batean oinarritutako ikastaroak eskaintzen baitituzte. Aitzitik, Seneko aplikazioaren funtzionamendua beste ikuspuntu batetik bideratu nahi izan da. Izan ere, fitxategiak eta ariketak partekatzeko metodoez baliatuz, kooperazioa eta elkarlana bultzatzen duen metodologia eskaini nahi da. Ondorioz, webgune-ko parte hartzaile guztiak erabiltzaileak izango dira, eta bakoitzak erabaki beharko du web aplikazioan izango duen inplikazio maila. Horregatik, erabiltzaileak ikasketa lanetara mugatu ez ezik, beste erabiltzaileen tutore lanak ere egiteko aukera izango dute; gainera, edozein kasutan posible izango da bi roll hauek momentu berean mantentzea.

Erabiltzaileen arteko elkarlan hau dinamikoki bideratzeko, beharrezkoa

izango da, besteak beste, erabiltzaile motak definitzea, fitxategiak eta ariketak partekatzeko metodoak definitzea eta fitxategien gainean burutu daitezkeen eragiketak zehaztea, honenbestez, web aplikazioan materiala kudeatzeko zein ariketak lantzeko guneak eskuragarri utzi beharko dira.

Metodologia berri hauen abantailen artean ondorengoak aurki daitezke: alde batetik, materialaren kudeaketa lanak erabiltzaile guztien artean banatzen direla, eta, ondorioz, lan karga ere banatu egiten dela, eta bestetik, eremu geografiko zabal batean lan egiteak kalitate eta aniztasun handiagoko materiala lortzea ahalbidetzen duela.

Aurretik azaldu dugun QG sistema web aplikazioaren motor gisa erabiltza, hizkuntza teknologiak ariketen sorkuntzaren eremuan aplikatzeko aukera izango da, eta, hainbestez, erabiltzaileek web aplikaziora igotzen dituzten testu fitxategien gainean galderak automatikoki sortzeko aukera egongo da. Dena den, sistema guztiz automatikoa izanik sortzen diren galderetan akatsak egotea aurreikusten denez, erabiltzaileek automatikoki sortutako material hau editatzeko -ezabatu eta zuzendu edo eduki berria sortzeko- aukera izango dute.

Honekin guztiarekin, erabiltzaileek hornitutako testuak erabiliko dira sistema informatikoaren datu iturri modura, eta hortik abiatuta, hizkuntzaren azterketa zein prozesamendurako teknikak erabiliz web bidez euskaraz idatzitako testuak lantzeko materiala sortuko da. Ondorioz, aplikazioaren osotasuna sistema erdiautomatiko gisa ikus daiteke, erabiltzaileak izango baitira sistemaren elementu pilare nagusiak, aplikazioak sortu dituen akatsak zuzenduz eta material berria gehituz.

Esaldi bakar batean laburtzearen, Senekoren helburua elkarlana eta metodologia kooperatiboa ahalbidetzen duen euskara ikasteko tresna interaktiboa izatea da.

Ondorengo kapituluetan, web aplikazio eta galdera sortzaile sistemen garapen osoa azaltzen da, gainera, proiektua garatzeko hartu beharreko erabakiak ikusi, lanaren deskonposaketa egiturak zehaztu, denbora plangintzak estimatu, software garapen prozesuak aukeratu, arriskuak analizatu, segurtasun kontuak lotu, erabilpen kasuak zein fluxu diagramak osatu, web aplikazio eta QG sistemen arteko komunikazioa zehaztu eta diseinu zein garapenean sortzen diren arazoei aurre hartu beharko diegu. Hala eta guztiz ere, azpimarratu nahi da, proiektu honetan inplementatu ez den atal bakarra *Etorkizuneko lanetan* kapituluan bildutako funtzionalitateak direla, beste funtzionalitate guztiak irteera produktuan inplementatu baitira.

Azkenik, aipatu, garatu den sistemak hiru ebaluaketa izan dituela: bata, hizkuntzalariekin prestatutako ebaluaketa, galdera sortzaile sistema ebaluatzea helburu izan zuena; bigarrena, web aplikazioaren ebaluaketa, unibertsi-tateko hainbat boluntarioekin egindakoa; eta azkena, hau ere web aplikazioa

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

ebaluatzeko, Martuteneko Arantzazuko Ama eskolan bederatzi/hamar urteko haurren ikasgela batean egindakoa. Ebaluaketa hauek guztiek, garatutako sistema ingurune errealago batean martxan jartzea izan dute helburu, eta etorkizuneko lanak zein urrats berriak emateko bideak argitzeko oso baliagarriak izan dira.

2.2 Proiektuaren abiapuntua

Aurkezten den proiektua ez da ezerezetik eratu den lan bat, 2010/2011 ikasurteko bigarren lauhilekoan zehar ikasleak proiektuaren zuzendariarekin batera egindako lan poltsa baten jarraipena du abiapuntu.

Lan poltsa hau IXA taldean¹ egikaritu zen 2011ko otsailetik hasi eta 2011ko irailera arte, guztira 320 ordutakoa izanik. Lan poltsa honen helburua ikaslea Lengoia Naturalaren Prozesamenduan eta Hizkuntza Teknologiaren arloko tekniketari zein tresnen diseinu eta erabilpenean trebatzea da. Betiere, ikasleak taldeak zuzendutako ikerkuntza proiektuetan parte hartzen duela.

Lan poltsak aurrera egin ahala eta formakuntzarako tartea zein beste lan batzuk egin ostean, proiektuaren zuzendariak ikerketa lerro berri bat eskaintzen du: Hizkuntzaren prozesamenduko teknikak irakaskuntza arlora aplikatzea. Honela, Itziar Aldabe Arregik eginiko *Automatic exercise generation based on corpora and natural language processing techniques*[14] izenburua duen tesian aurkeztu zen atal baten garapenarekin eta galdera sorkuntzaren inguruan[13] emandako lehen urratsekin jarraitzeko ideia jorratzen da. Lan poltsa honetan zehar edozein testutik abiatuta galdera irekiak sortzen dituen aplikazioaren lehen bertsioa garatzen da. Lan poltsa bukatzean, proiektuaren zuzendari izango den eta ikaslearen artean buruturiko lan honi jarraipena ematea erabakitzen da karrera bukaerako proiektu gisa.

Buruturiko bi lan hauen arteko bereizmen zuzena argitzeko idatzi da azpi kapitulu hau. Bi lanak bata bestearekin erlazionatuta daudelako, eta ondorioz argi utzi nahi delako lan poltsan egindako lanaren norainokoa eta nola lotzen den karrera bukaerako proiektuarekin.

¹IXA taldeak hogeita lau urte daramatza Lengoia Naturalaren Prozesamendua eta Hizkuntza-Teknologia arloetan lanean. Arlo zabal horien barruan, euskararen gaineko ikerketa aplikatua da taldearen xede nagusia, baina beste hizkuntzen gaineko ikerketa zein produktu komertzialen garapena ere sartzen dira buruturiko eginkizunen artean. Lan horretan diziplinarteko taldea osatzen du, informatikariak eta linguisten artean batez ere. Azken helburua euskararako hain beharrezko eta onuragarri izan daitezkeen sistema (erdi)automatikoak egitea edo egiten laguntzea da, teknologia berrien eragina hizkuntzan ahalik eta onena izan dadin.

Galdera sortzailearen azalpen laburra

Lan poltsan garatutako lanaren helburua esaldi mailan galderak masiboki sortzeko gai den sistema garatzea izan zen. Irakurlea koka dadin, tresna honen azalpen laburra egitea beharrezkoa dela ikusi da, proiektuaren hasieratik galderen sorkuntzaren inguruan arituko baikara.

Oro har, galderak automatikoki sortzeko tresnek jarraitzen duten prozedura lau etapatan banatzen da:

1. Sarrerako testua tratatu
2. Galderen emaitza izango diren hautagaiak aukeratu
3. Galdetzaileak aukeratu
4. Galdera eraiki

Lehen hiru pausoak: sarrerako testua tratatu, edukia aztertu eta galdetzailea identifikatu, informazio linguistikoan oinarrituz burutzen dira. Informazio linguistiko hau Natural Language Processing (NLP) tresnak erabiltzaz lortzen da, esaterako: analizatzaile sintaktikoak, entitateak hautemateko analizatzaileak ... Aipatu beharra dago, analisi hauek burutzerako garaian IXA taldean garatutako tresnak erabiltzen direla, hala nola: Ixati[34], euskararen analisi sintaktikoa egiteko tresna eta Eihera[32], pertsonak, tokiak eta erakundeak sailkatzeko tresna. Laugarren pausoa aldiz, galderazko perpausa eraikitzea, eraldaketa erregela zein patroietan oinarrituta burutzen da. Eraldaketa erregela zein patroik hauek jatorrizko esaldiaren hitzak berrantolatzen dituzte galderazko perpausak osatzeko.

2.2.1 Proiektura egokitzeko egindako aldaketak

Lan poltsan galderak automatikoki sortzeko (QG) sistema garatu bada ere, proiektuaren helburua desberdina da. Aurreko ataletan aipatu den modura, proiektuaren helburua euskaraz idatzitako testuak lantzeko aukera ematen duen aplikazioa garatzea da, hala ere, proiektu honetan bi sistemak elkarlanean arituko dira, web aplikazioaren motorra lan poltsan ikasleak garatutako QG sistema izango baita.

Dena den, bi lanen arteko elkarrekintza ahalbidetzeko moldaketa eta egokitzapen batzuk egitea ezinbestekoa izango da galdera sortzailean. Horregatik, lehen bertsioaren gainean lan egitea beharrezkoa izango da. Aldaketa hauek zehatzago diseinuari dagokion kapituluan azalduko badira ere, helburu bikoitza izango dute: alde batetik, web aplikazioa eta galdera sortzaile

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

sistemaren arteko komunikazioa ahalbidetzea, honetarako, sarrera eta irteerak egokitzea eta markatze lengoaietan oinarritutako formatu berriak zehaztea beharrezkoa izango da. Bestetik, galdera sorkuntza prozesua sendotzea, honetarako, aurreko atalean ikusitako faseak hurbilpen linguistiko berriekin osatuz.

2.3 Ataza eta azpiatazen zerrenda

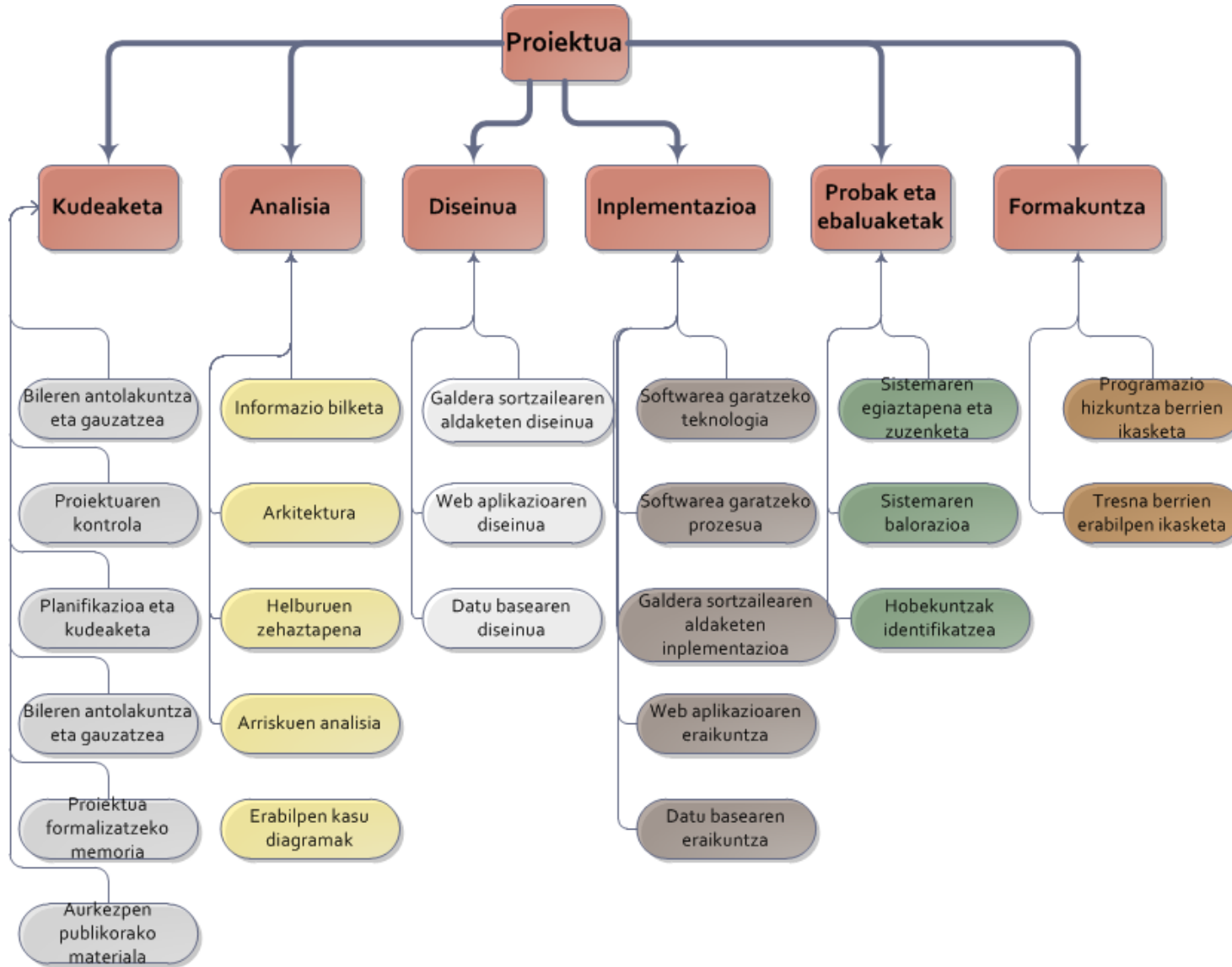
Aurkezten den proiektuan sei fase nagusi bereizten dira. Fase hauek aplikazioa garatzeko kontuan hartu diren atazak biltzen dituzte eta proiektuaren bizi zikloa islatzen dute.

- Proiektuaren kudeaketa.
 - Lana kudeatzeko eta planifikatzeko azpiatazen multzoa.
 - * Bileren antolakuntza eta gauzatzea, proiektuaren kontrola, planifikazioa eta kudeaketa, proiektuaren formalizaziorako memoria eta aurkezpen publikorako materiala.
- Analisia.
 - Proiektua garatzeko beharrezko informazioaren bilketa edota zirikusia duten aplikazioen analisia.
 - Aplikazioaren arkitektura.
 - Helburuen zehaztapena.
 - * Proiektuaren ezaugarriak finkatzea.
 - Arriskuen analisia.
 - Erabilpen kasuen diagrama.
- Diseinua.
 - Galdera sortzailearen diseinua.
 - Web aplikazioaren diseinua.
 - * Datu basearen diseinua.
 - * Web aplikazioaren diseinua.

- Implementazioa.
 - Analizatu eta diseinatutako sistemaren implementazioa burutzea.
 - * Softwarea garatzeko teknologien aukeraketa.
 - * Softwarea garatzeko prozesua.
 - * Galdera sortzailearen implementazioa.
 - * Datu basearen eraikuntza.
 - * Web aplikazioaren eraikuntza.
- Probak eta ebaluaketak.
 - Garatu den sistema egiaztatu, zuzendu, baloratu eta hobekuntzak identifikatu.
- Formakuntza
 - Proiektua garatu ahal izateko beharrezko gaitasunak identifikatu eta lortzea.
 - * Teknologia berriak barneratzea eta baliagarriak izan daitezkeen tresnen erabilera ikastea.

2.4 Lanaren Deskonposaketa Egitura (LDE diagrama)

Lanaren Deskonposaketa Egitura diagraman proiektuaren norainokoak finkatzen dira era multzokatu, hierarkiko eta egituratuan. Honela, burutu beharreko lan karga atazetan banatzen da era grafikoan. Ataza eta azpiataza hauek aurreko atalean azaldutakoak dira.



Irudia 2.1: Lanaren Deskonposaketa Egitura (LDE diagrama).

2.5 Proiektuaren plangintza

Atal honetan aurkezten den proiektuaren garapena azalduko da denboran zehar, proiektuaren parte diren ataza eta azpiataza guztiak noiz garatuko diren finkatuz.

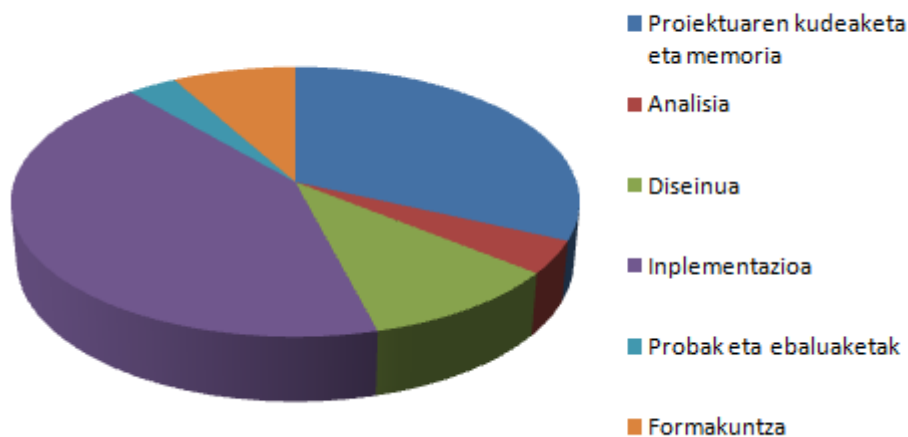
2.5.1 Estimazioa

Ondorengo taulan ataza eta azpiataza guztientzat egindako denbora estimazioak aurkezten dira. Proiektuaren kudeaketaren ondorioetan estimatutako denboren desbideraketak aztertu ahalko dira.

Atazak eta azpiatazak	Estimazioak
Proiektuaren kudeaketa eta memoria	138 ordu
Bileren antolakuntza eta gauzatzea	40 ordu
Proiektuaren kontrola	5 ordu
Planifikazioa eta kudeaketa	8 ordu
Proiektuaren formalizaziorako memoria	75 ordu
Aurkezpen publikorako materiala eta prestaketa	10 ordu
Analisia	17 ordu
Informazioaren bilketa eta aplikazioen analisia	6 ordu
Aplikazioaren arkitektura	3 ordu
Helburuen zehaztapena	4 ordu
Arriskuen analisia	1 ordu
Erabilpen kasuen diagramak	3 ordu
Diseinua	44 ordu
Galdera sortzailearen diseinua	18 ordu
Datu basearen diseinua	3 ordu
Web aplikazioaren diseinua	23 ordu
Inplementazioa	185 ordu
Software garatzeko teknologien aukeraketa	5 ordu
Softwarea garatzeko prozesua definitu	3 ordu
Galdera sortzailearen implementazioa	25 ordu
Datu basearen eraikuntza	2 ordu
Web aplikazioaren eraikuntza	150 ordu
Probak eta ebaluaketak	14 ordu
Formakuntza	35 ordu
Ataza eta azpiataza guztien batura	433 ordu

Taula 2.1: Proiektuaren denbora plangintza.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa



Irudia 2.2: Proiektuaren denbora plangintza.

Irudian ikus daitekeen modura, implementazioa da karga handiena duen ataza. Hala ere, diseinu eta analisiaren atazek ere karga handia dute proiektuaren osotasunean. Izan ere, egin beharreko lanaren definizioa ezinbestekoa da implementazioari ekin aurretik.

Formakuntzaren atalari esleitutako karga ere nabarmena da, web aplikazioa garatzerako garaian batez ere, teknologia ezezagunak erabiltzea aurreikusten delako.

Proiektuaren kudeaketari ere garrantzia eman nahi izan zaio, karrera bukaerako proiektu bat izanik ataza honetan murrizketarik ez egitekotan.

2.5.2 Gantt diagrama

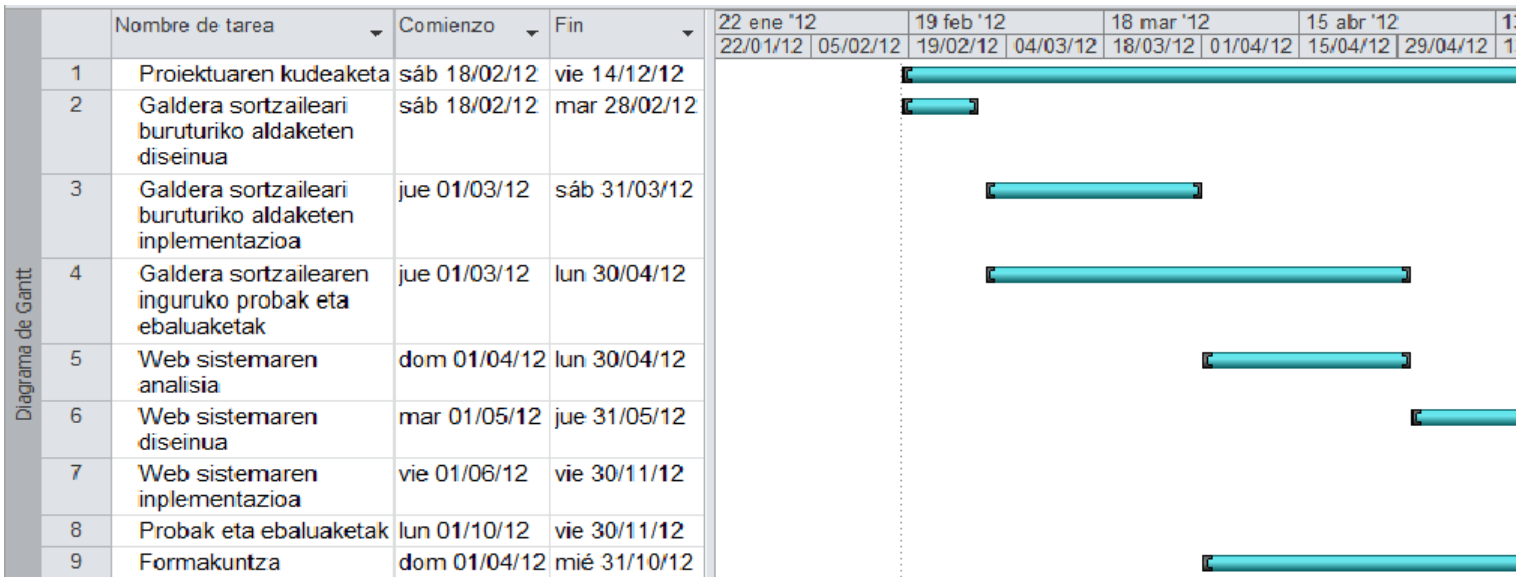
Atal honetan, planifikatutako ataza eta azpiatazen egutegiaren adierazpide grafikoa egingo da. Gantt diagramaz baliatuta, proiektuak izango duen fase bakoitzaren denbora estimazioaren hasiera eta bukaera epeak finkatuko dira.

Ahal den neurrian estimatutako denborak errespetatzen ahalegindu behar da baina, softwarea garatzeko prozesuen azpiatalean aipatzen den modura (5.1 azpiatala), garapen prototipatu inkremental eta ebolutiboa oinarritzat hartuz, gerta liteke erabiltzaileen feedbacka jasotzean diseinu edota betekizunen zehaztapenak aldatzea. Honela jazotzen bada, birplanifikazioak edota desbideratzeak dokumentatuko dira (7. kapitulua). Izan ere, aplikazioaren bizi zikloaren hasierako unetik prototipo bat egongo da atzigarri erabiltzai-

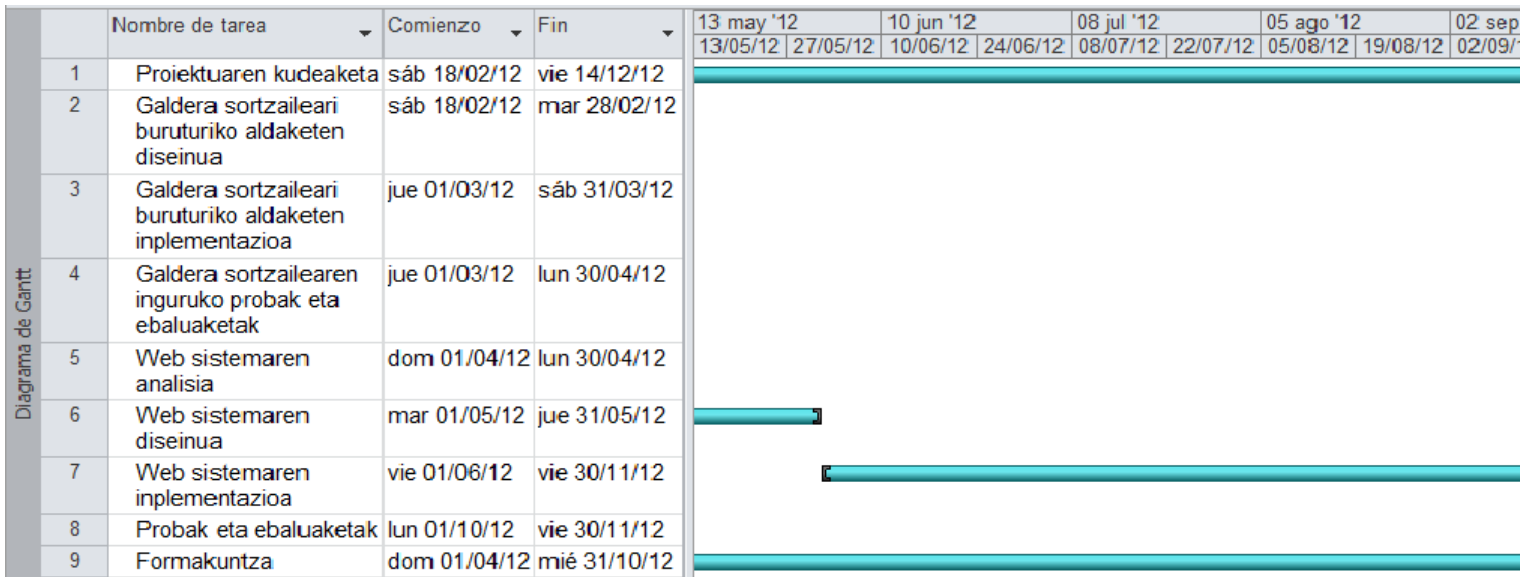
Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

leentzat sarean. Honenbestez, prototipo hau erabiliz software garapenean zehar erabiltzaileen balorazioa edota iritzia jasoko da, eta, era honetan, funtzionalitateak garatu bezain azkar ebaluatzea ahalbidetuko da. Softwarea garatzeko prozesu honen bitartez arazo, akats edota funtzionalitate aldaketen hautemate goiztiarra burutzea espero da.

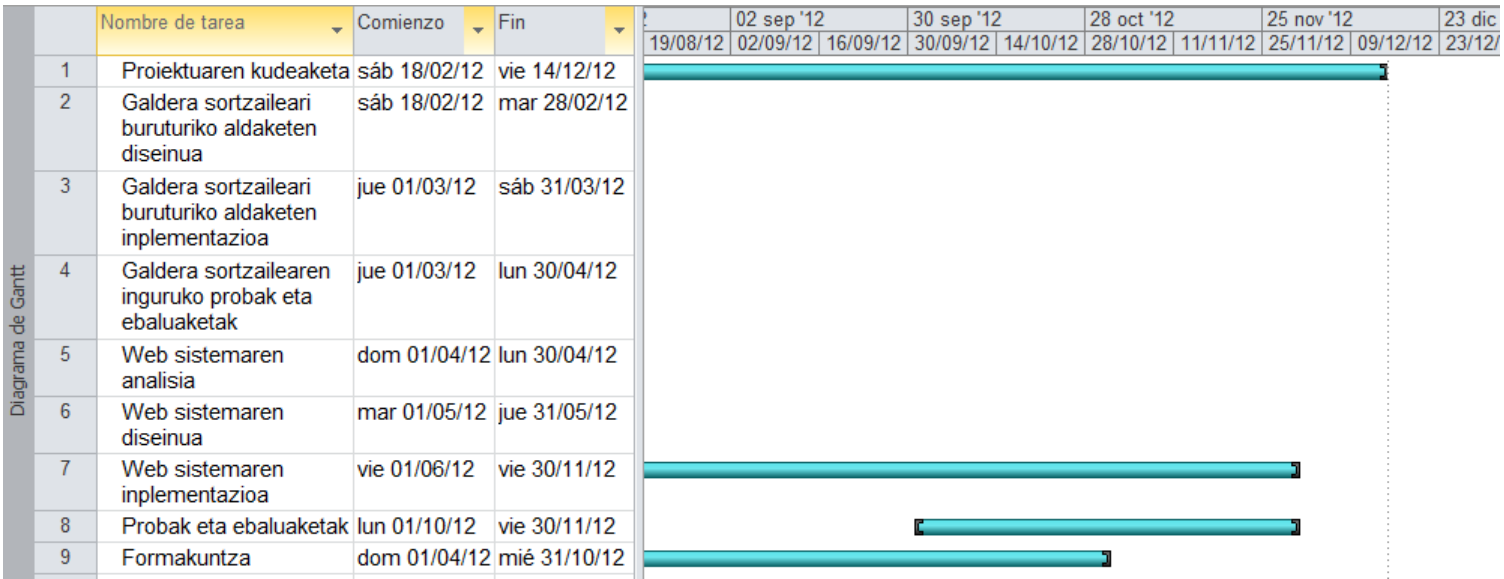
Proiektua 2012 urtean zehar garatuko da, 2012ko otsailean hasiz eta 2012ko abenduan amaitzea espero da.



Irudia 2.3: Gantt diagrama.



Irudia 2.4: Gantt diagrama.



Irudia 2.5: Gantt diagrama.

Irudian ikus daitekeen modura, proiektuaren bizi zikloa hamaika hilabetetan zehar banatu da. Bizi zikloaren hasierako faseetan galdera sortzailerengan ipiniko dugu arreta, eta ondoren, web aplikazioa garatzea izango da helburu nagusia. Proiektuaren kudeaketa eta formazioari dagokionez, bizi zikloaren zati handi batean agertzea espero da. Izan ere, proiektuaren kudeaketaren zati dira bileren burutzea eta dokumentazioaren zirriborroak idaztea besteak beste eta ataza hauek proiektuaren garapeneko fase guztietan zehar hedatuko dira.

Proiektua erabat garatzeko estimazioa 433 ordutakoa izan da eta hamaika hilabetetan zehar banatzen denez, hilabete bakoitzean berrogei ordu inguruko lan karga estimatzen da. Honenbestez, lan karga astunegia ez izanik, suertatu daitezkeen arazoen aurrean malgutasuna izatea espero da arazoei aurre egiteko.

2.6 Lan metodologia

Proiektua zentzuz garatu ahal izateko, lehendabiziko faseetan modu eta baldintza egokiak zehaztea beharrezkoa da.

Bilerei dagokienez, astero edota gehienez jota hamabost egunero beharrezkoak izango direla estimatzen da. Honetaz gain, ezohiko bilerak egotea espero da eta hauek beharraren arabera finkatuko dira bizi zikloan zehar. Bileren iraupenari dagokionez ordu bete edota ordu eta erdi ingurukoak izatea estimatzen da, betiere jorratzen diren gaien menpe egonik. Gainera, bilera guztien ostean ikasleak korreoz bidali beharko dio proiektuaren zuzendariari bilerako kontzeptu garrantzitsuenak laburbiltzen dituen akta.

Ikasleak proiektuan lanean jardungo duen denborari dagokionez eguneko hiru ordu izatea espero da, dena den, desbideraketak egotea espero da beste irakasgaietako lan kargak baldintzatuta, gehienbat azterketa garaian.

Proiektuaren kudeaketari dagokionez, hainbat fase xedatu dira lana astunegia egin ez dadin. Fase bakoitzaren emaitzak zirriborro batean idatziko dira eta behar izanez gero, bileretan proiektuko zuzendariari utziko zaizkio honek azter ditzan. Honela lan eginez gero, dokumentazioaren lan karga bizi ziklo osoan zehar banatzea espero da.

Proiektua garatzeko gidalerro nagusiak ondorengoak dira:

- Lehenengo fasean, proiektuaren helburu dokumentatuen atala idatzi eta aurrekarien aztertzeari ekingo zaio. Era honetan, lortu nahi diren helburuak zehaztuko dira, eta, gainera, proiektuaren garapenerako lagungarria izango diren beste erabaki batzuk ere hartuko dira, esaterako: lan metodologia edota software prozesua.

- Bigarren fasean, proiektuaren abiapuntu den galdera sortzailea programa birmoldatuko da.
- Hirugarren fasean, proiektuaren abiapuntua den galdera sortzailearekin elkarrekintza izango duen web aplikazioa garatzea da helburua.
 - Web aplikazioaren analisia, diseinua, implementazioa eta ebaluaketa burutuko da.
- Laugarren fasean, bizi ziklo guztian zehar idatzitako zirriborroak formalizatu egingo dira karrera bukaerako proiektuaren txostena osa dezaten.

2.7 Bideragarritasuna

Aurkezten den proiektua bideragarria izatea ahalbidetzen duten baldintzarik garrantzitsuenak ondokoak dira:

- Proiektuaren helburuak aztertu, finkatu, eztabaidatu eta onetsi dira proiektuaren bizi zikloaren hasierako unetik.
- Proiektuaren plangintza eta kudeaketa aproposa bermatzen duen lan metodologia definitu da.
- Proiektuaren garapenerako lagungarria izango den software prozesua definitu da.
- Proiektua garatzeko software librea zein MDNAA ² bidez atzigarri dagoen softwarea erabili da. Era honetan, ez da baliabide ekonomikoren beharrik izango proiektua bere osotasunean garatzeko.

2.8 Arriskuen Analisia

Proiektuaren egikaritzapenean zehar hainbat arrisku sor daitezke. Arrisku hauek softwarea garatzearen ondorio dira eta arrisku hauen aurreikuspena burutzea oso garrantzitsua gerta daiteke. Arriskuen analisiarekin hauek sor ditzaketen kalteak ekidin edota murriztea da helburua, honetarako beharrezkoa izango da arriskuak identifikatzea eta dagozkien kontingentzia planak zehaztea.

²Microsoft Developer Network Academic Alliance EHUko informatika fakultateak eskaintzen duen baliabide bat da. Zerbitzu honen bitartez EHUk informatika fakultateko ikasleek Microsofteko softwarea erabil dezaten ahalbidetzen du. Oharra: 2012ko otsailetik aurrera zerbitzu hau DreamSpark bezela ezagutarazten da.

2.8.1 Arriskuen zerrenda

Proiektuaren inguruan aurreikusten diren arriskuen zerrenda ondorengoa da:

1. Zehaztutako plangintza betearazteko gai ez izatea irakasgaiek pentsatutakoa baino lan karga astunagoa suposatzen badute.
2. Softwarea garatzeko zehazturiko prozesua dela eta, erabiltzaileekiko elkarrekintza astunegia gertatzea. Funtzionalitate batzuk aldatzea espero da erabiltzaileen feedbacka jaso ahala baina kopurua altuegia izanez gero zehaztutako plangintza betearazteko arazoak aurreikusten dira.
3. Proiektuaren inplementazioarekin erlazionatutako datuen galera.
4. Proiektua martxan egoteko beharrezkoak diren zerbitzarietan arazoak egotea, hala nola, web zerbitzarian edota datu basearen zerbitzarian.

2.8.2 Arriskuen kontingentzia plana

Aurreko arriskuen kaltea deuseztatu edota murrizteko asmoz ondoko kontingentzia plana diseinatu da:

1. Proiektuari dagokion plangintza era malguan diseinatu da, lan astunegia esleitu gabe. Era honetan atzerapenen aurrean lan orduen kopurua areagotzea posible izango da. Gainera, beharrezkoa suertatzen bada, proiektua birplanifikatuko litzateke entregatze datak atzeratuz.
2. Erabiltzaileek emandako gida lerroak kontsideratuko dira baina zehazturiko proiektuarekiko desbideratze handiegia suposatzen badute baztertu egingo dira. Nahiz eta denboraldi baterako baztertuak izan dokumentatuta gordeko dira, eta, era honetan, posible izango da etorkizuneko lan gisa bere inplementazioa kontsideratzea beste garapen iterazio bat egitea erabakitzen bada.
3. Datuen galera ekiditeko backup politikak erabiliko dira. Erabiliko diren zerbitzarietan politika hauek funtzionamenduan daude, ordenagailu pertsonaletan aldiz, politika berriak ³ ezarriko dira.
4. Zerbitzarietan arazoak egoten badira, gehienez ere aste bateko epean konponduko direla estimatzen da. Horregatik, garapenarekin jarraitzeko ordenagailu lokalean instalatutako web zerbitzaria eta datu base

³4.7.1 azpiatalean aurki daiteke erabilitako backup politika

zerbitzaria erabili beharko lirateke. Honen ostean, zerbitzarietako arazoak konpontzean, lokalean garatutako funtzionalitateak irauli beharko lirateke hauetara. Dena den, tarte honetan webguneko atzipena bermatu gabea egongo litzateke. Ondorioz, arazoa luzatuko balitz, sareko beste zerbitzari batzuk erabiltzera behartuak egongo ginateke. Betiere, lan karga kontuan izanik eta denbora asko galdu gabe, zerbitzaria aldatzeak lan handia suposatzen baitu.

3 Kapituluia

Analisia

Gaien Aurkibidea

3.1 Erabiltzaileen beharren eta antzeko tresnen azterketa	22
3.1.1 Ondorioak	24
3.2 Aplikazioaren arkitektura	26
3.3 Erabilpen kasuen analisia	30
3.3.1 Erabiltzaile kautotu gabeen erabilpen kasuen analisia	31
3.3.2 Erabiltzaile kautotuen erabilpen kasuen analisia	32
3.3.3 Administratzailearen erabilpen kasuen analisia	33
3.4 Software erabakien eta teknologiaren analisia	34
3.4.1 Erabilitako azpiegitura, teknologia eta baliabideak	34
3.4.2 IXA taldean garatutako tresnen eta baliabideen erabilera	40
3.4.3 Frameworken inguruko erabakia	41
3.4.4 Txostena LaTeX bidez garatzeko erabakia	42

Atal honetan sistemaren analisi osoa egingo da. Sistema analizatzerakoan proiektuaren helburuak eztabaidatuko dira eta zehaztasun osoz eginbeharrak zerrendatuko dira.

3.1 Erabiltzaileen beharren eta antzeko tresnen azterketa

Sistemaren analisi sakon bat egin aurretik, erabiltzaileen beharren eta antzeko tresnen analisia egin behar da. Atal hau burutzea oso lagungarria izango da ondoren etorriko diren faseak garatzeko edota hartu behar diren lanak eta erabakiak errazteko. Esaterako, erabilpen kasuei, datu baseari eta web interfazeari dagozkien erabakiak bideratzen hasteko oso aproposa izango da.

Atal honen helburua garatu behar den aplikazioaren antza duten beste aplikazioen analisia egitea da. Ikasketa sistemari dagokionez, *hiru.com* aplikazioa aztertzea erabaki da, oso interesgarria den eta toki askotan erabiltzen den -euskaltegietan esaterako- aplikazioa baita. Gainera, garatutako aplikazioa ariketak eskaintzera bakarrik mugatzen ez den heinean, hau da, testu fitxategietatik abiatuta galderak automatikoki sortzeko eta sortutako material hau kudeatzeko aukera ere ematen duenez, beste aplikazio batzuk aztertzea beharrezkoa izan da, kasu: *ikasbil* eta *ArikIturri*.

Analizatuko diren web aplikazioei dagokienez -*hiru.com* eta *ikasbil*- aipatu beharra dago, oro har, eduki estatikoz osatutako ikastaroak eskaintzen dituztela. Alabaina, proiektu honetan *ArikIturri* tresnak jarraitzen duen filosofiaren tankerako metodologia jarraituko da; hau da, ikasketarako material itxia eskaintzen duten sistemen alternatiba gisa. Izan ere, metodologia honetan erabiltzaileak izango dira sistemaren antzezle nagusiak: fitxategiak igotzeaz, automatikoki sortutako materiala kudeatzeaz, osatzeaz eta parte-katzeaz arduratuko direlako.

Hiru.com

Hiru.com hezkuntza alorreko ataria sortzeko lehen ahalegina da eta bere helburua hezkuntzako zerbitzu publikoak internetetik ematea da, bere edukia osotasunean doan eskainiz. Hiru.com Etengabeko Ikaskuntza atari; Eusko Jaurlaritzaren Hezkuntza, Unibertsitate eta Ikerketa Sailaren ekimena da, eta bere asmoen artean ondorengoak daude, besteak beste: Euskal hiritarren prestakuntzarako erreferente bilakatzea eta hiritarrei laguntzea -euren jarrera aktibo eta parte hartzailea sendotuz-.

Atari hau sekuentzia logiko baten lehenengo urratsa da. Izan ere, Euskal hiritarren eguneroko bizitzan IKTak integratzeko asmoz, 1999an Eusko

Jaurlaritzak Euskadi 2000 Hiru egitasmoari ekin zionetik, aurrera eramaten ari den ekintzen artean kokatu beharra dago. Une honetan, eta hiru urteko esperientziaren ondoren, jakin badakigu atari honek erabiltzaileen arreta jaso duela, eta herritarren, zein familien formakuntzan lagungarri izaten ari dela.

2003an bere ibilbidea hasi zuenetik, Hiru.com ataria etengabe aberastu eta eguneratu da, gero eta eduki gehiago eskainiz, betiere gizarteak dituen premia eta interesei erantzuteko asmoarekin.

Webgune honetan bi atal nagusi bereiz daitezke: E-lea eta ikasten.net.

E-lea

- Atzerriko hizkuntzak ikasteko baliabidea, E-lea ikastaroekin lau hizkuntza ikas daitezke (ingeleza, frantsesa, alemana eta italiera).
- on-line ikastaroak eskaintzen ditu, era honetan norberak bere erritmoa mantentzeko aukera emanik.

Ikasten.net



Irudia 3.1: ikasten.net logotipoa.

Ikasten.net euskara ikasteko baliabide didaktikoa da, munduko edozein tokitan dauden pertsona helduei euskara helarazteko metodologiaz diseinaturata, interfaze atsegin eta informolarekin. Ataletan antolatutako hirurogeita hamar unitate didaktikoz osaturiko ikastaroa, eta errepasorako hamabi unitate eta hainbat ekintza dakartza: audioak, bideoak eta hainbat motatako ariketak -hutsuneak bete, lotzea edo sailkatzea, aukera anitza, irakurmena, teoria, idazmena, sailkatu gabea, erantzuna hautatu, jarduera gehigarria, hiztegi-fitxak, grabaketa, ... -.

2011ko urritik aurrera jarduerak zuzentzeko zerbitzua dago eskuragarri. Izan ere, badira automatikoki zuzendu ezin daitezkeen jarduerak -idazlanak, ahozko grabaketak, e.a.-. Horiek irakasle batek zuzentzeko aukera du orain ikasleak.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Ikasbil

Ikasbil ataria euskara ikasten eta irakasten ari diren herritarrei zuzendutako HABEren¹ web ataria da. Bertan euskara ikasteko lagungarri izango den eduki eta material mota ugari dago -eta maiztasun handiz eguneratzen da- esaterako: entzungaiak, ariketak, argitalpenak, ... Baliabide hauez gain, erabiltzaileentzat lagungarri diren beste hainbat zerbitzu ere eskaintzen dira: euskaltegiei buruzko informazioa, euskaltegietak ikastaroak, egindako lanaren erregistroa, HABEko liburutegia eta albisteak, ...

Gainera, atari honen azken bertsioak erabiltzaileari Ikasbilen sare sozialean parte hartzeko aukera ematen dio, eta, era honetan, elkarlana bultzatu euskara ikasteko interesa duten pertsona guztiak harremanetan jarriz.

ArikIturri

ArikIturri [12] euskararako ariketa desberdinak sortzeko gai den eta arkitektura irekia duen sistema da. Ariketa sortzaile honen informazio iturria XML markatze lengoiaian adierazita dauden linguistikoki analizatutako corpus errealak dira eta sortutako ariketak ere XML markatze lengoiaz adierazten dira.

Sistema honek automatikoki ariketak sortu ahal izateko bi hizkuntza baliabide desberdinetan oinarritzen da: alde batetik, Lengoia Naturalen Prozesamendurako tresnetan, eta, bestetik, ariketa sorkuntzarako informazio linguistikoko espezifikoa.

Oro har, tresna hau hutsuneak bete, hitz elkarketak, aukera anitza eta akatsak zuzendu motako ariketak sortzeko gai da.

3.1.1 Ondorioak

Aurreko sistema guztiak aztertu ondoren, gure sistemak bildu behar dituen ezaugarriak zehaztu² genituen.

1. Ariketak euskaraz bakarrik eskainiko dira, bai garatu den galdera sortzailea, baita galdera sortzailea oinarritzen den tresna linguistikoko guztiak euskarara mugatzen baitira.
2. Eskainitako ariketak galderak erantzun motakoak izango dira, proiektuaren helburuetako bat galdera sortzailea garatzea delako.

¹<http://www.habe.euskadi.net>

²Ondoren zerrendatzen diren ezaugarri guztiak garatu den sisteman inplementatu dira. Bildutako, baina inplementatu gabeko, ezaugarriak etorkizuneko lanetan azaltzen dira.

3. Irakaskuntza/ikaskuntza metodologia kooperatiboa bultzatu nahi da. Honetarako, euskarri gisa fitxategi eta ariketen partekatzea erabiliko da.
4. Interfazeak erabiltzeko erraza behar du izan, adin guztietako erabiltzaileak izatea espero baita. Honenbestez, erabiltzaileak sortzeko; fitxategiak kudeatzeko -automatikoki sortutako galderak editatzeko, zuzentzeko edo ezabatzeko zein galdera berriak gehitzeko-, igotzeko eta hauen gainean bilaketak egiteko eta ariketak automatikoki sortzeko; ariketen zailtasun mailak ikusteko eta esleitzeko; erabiltzaile datuak aldatzeko; pasahitza berreskuratzeko; ariketak lantzeko; ... interfazeak erabilerraza izan behar du.
5. Erabiltzaileen arteko elkarlana sustatzea garrantzitsua da. Elkarlan honek giro hobea izatea izango du ondorio, honetarako, erabiltzaileen artean mezuak partekatzeko txata erabilgarri egongo da.
6. Erabiltzaileen ikuspuntutik fitxategi edota ariketen inguruan balorazioak eta iritziak uzteko baliabideak eskaini behar dira.
7. Erabiltzaileari zereginak erraztearren, ariketak lantzeko garaian laguntzazko baliabide batzuk eskaintzea zilegi da, esaterako: euskal hiztegiak eta itzultzaile automatikoak. Honetaz gain, galderen emaitzak asmatzen lagunduko duen baliabidea ere eskainiko da tresna erabilgarrien artean. Gainera, erabiltzaile bakoitzak ariketak nahi haina aldiz burutu eta zuzendu ahal izango ditu, berriro hasi, aurrekora itzuli, ... eta horrelako mugarik ezarri gabe.
8. Irakaskuntza gidak ez du amaiera zehatz bat izan behar, erabiltzaileak erabakiko du sistema noiz uzten duen. Ondorioz, unitate didaktikorik ez da izango aplikazioan; ostean, erabiltzaileek hornituko dute aplikazioa hauek gogoko dituzten testuak igoz. Era honetan, formakuntzarako materiala mugagabea -erabiltzaileek igotzen dituzten fitxategi kopuruari proportzionala- izango da.
9. Ariketak lantzeko garaian ez da inolako ordenarik edo murriztapenik ezarriko, edozein erabiltzailek edozein unetan nahi duen fitxategia erabiltzeko gai izango da eta, ez hori bakarrik, fitxategi horren inguruan sortu diren galderak ere edozein ordenatan lantzeko aukera izango du.

Aurretik aipatu den modura, jasotako ideia hauek guztiak egokiak direla uste da eta metodologia berrietako baliabideak eskaintzen dituztela. Ondorioz, erabilpen kasuak analizatzen eta diseinatzen direnean zerrendatutako puntu guztiak islatuko dira.

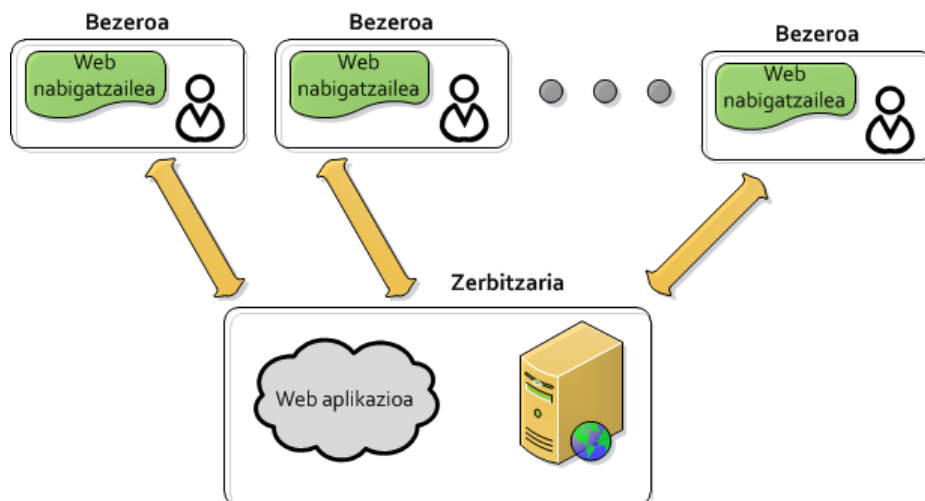
Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

3.2 Aplikazioaren arkitektura

Malgutasun, moldagarritasun eta hedagarritasun printzipioei jarraituz, aplikazioa software garapenean erabiltzen diren arkitektura eskemak oinarritzat hartuz garatu da. Arkitektura eskemei dagokienez, zehazki, bezero zerbitzari eskema, geruzetan antolatutako programazio eredua eta Model-View-Controller (MVC) eredua kontuan izan dira. Web zerbitzuen arteko komunikazioa, hurrengo kapituluetan sakonago azalduko bada ere, XML dokumentuen bitartez egiten da. Gainera, IXA taldeak garatutako hainbat tresna eta baliabide web zerbitzu autonomo gisa berrerabili dira.

Bezero zerbitzari eredua

Oro har, web aplikazioak bezero zerbitzari arkitekturetan oinarrituta egon ohi dira; izan ere, eredu edo arkitektura honek aplikazioen ahalmena zerbitzariaren eta bezeroen artean banatzeko aukera ematen du; eta, ondorioz, lana eraginkortasun handiagoz burutzen da. Arkitektura honetan zerbitzariak web aplikazioa eskaintzen du eta erabiltzaileek beren makinetatik webgunea atzitzen dute modu konkurrentean. Honetarako, nahikoa izango da zerbitzariak HTTP zerbitzua eskaintzea eta bezeroek, aldiz, HTTP eskakizunak interpretatuko dituen softwarea izatea, hau da, web nabigatzailea.



Irudia 3.2: Bezero zerbitzari ereduen eskema.

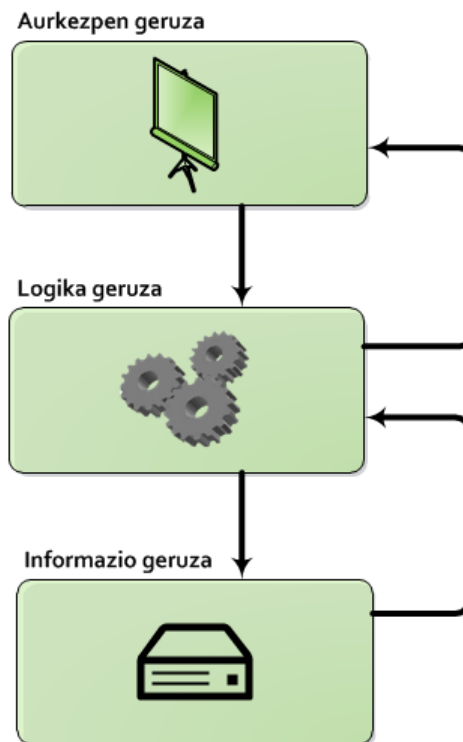
Geruzetan banatutako programazio eredua

Garatutako aplikazioaren osagai desberdinak logikoki banatzeko asmoz, geruzetan³ banatutako programazio eredua erabili da. Programazio teknika honen bitartez aplikazioa garatzeko kontrol handiagoa eskuratzen du programatzaileak. Geruzetan banatutako programazio eredu zabalduena hiru geruzetan oinarritzen da:

1. Aurkezpen geruza (Client tier): Bezeroak jasotzen duen edukia biltzen du, hau da, aplikazioaren interfazea.
2. Logika geruza (Logic tier): Web aplikazioaren edukia implementatzen duen geruza.
3. Informazio geruza (Data tier): Informazioa atzitzeaz eta biltegitratzeaz arduratzen den geruza.

Programazio eredu honek eskaintzen duen abantailarik nagusiena ondorengoa da: kodea logikoki banatua dagoela, eta, ondorioz, garapen eta moldaketa lanak burutzeko beharrezko denbora laburtzen dela osagaien gaineko atzipena zuzenagoa, garbiagoa eta azkarragoa delako.

³Maila desberdinez ere hitz egin genezake, mailak fisikoki banaturik dauden geruzak baino ez baitira



Irudia 3.3: Geruzetan banatutako programazio ereduaren eskema.

MVC eredua

MVC eredua (Model, View eta Controller) software arkitektura eredu bat da, interfazea, logika eta informazioa hiru osagai desberdinetan banatzen dituenena. Arkitektura honen erabilera oso hedatua dago web aplikazioetan, atal logikoak banatuta mantentzeak kode aldaketak egiteko erraztasunak ematen baititu. Arkitektura honen osagaiak ondorengoak dira:

- Model

Web aplikazioaren logika geruzak eta informazio geruzak osatua, webgunearen programazio guztia biltzen eta informazioa atzitzeaz arduratzen den osagaia.

- View

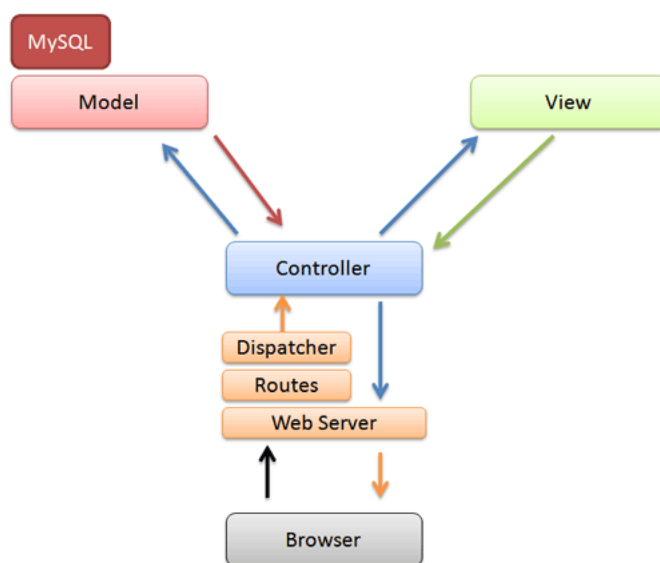
Bezeroak jasotzen duen interfazea biltzen duen osagaia.

- Controller

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Atazak hasieratzen dituzten abiarazleak dira hauek, erabiltzaileen elkarrekintzaren ondorioz sorturiko ekintzak jasotzeaz eta emaitza gisa ataza bat martxan jartzeaz arduratzen den osagaia.

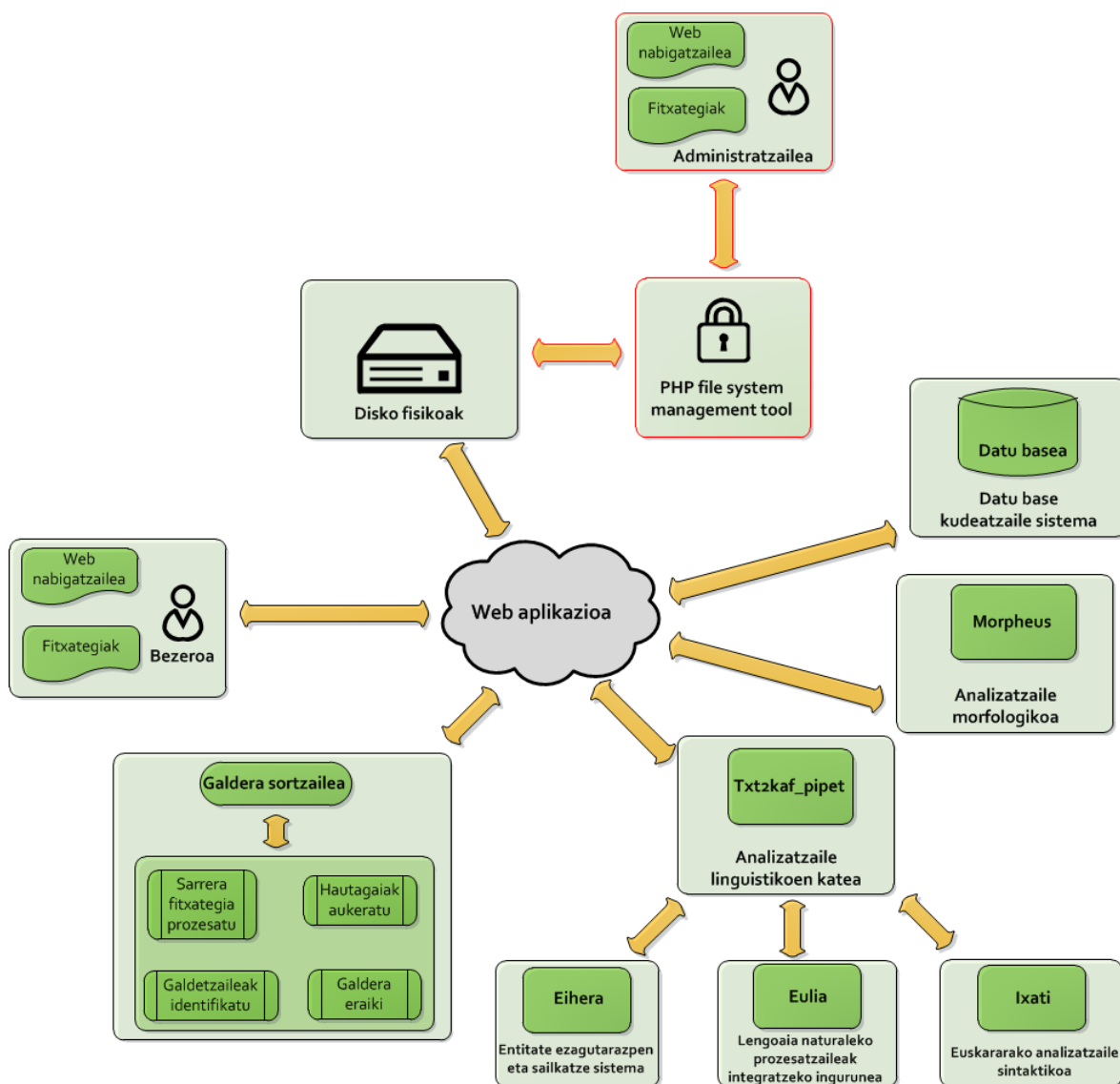
Arkitektura eredu honek hiru geruzetan antolatutako programazio ereduarekin antza badu ere, ez da gauza bera. Izan ere, MVC ereduak elkarrekintzei erantzuteko eskema bat zehazten duen bitartean, geruzetan banatutako programazio ereduak kodearen ulergarritasuna areagotzeko banaketa logiko egokia zehaztea du helburu.



Irudia 3.4: MVC ereduaren eskema.

Azpiegitura

Aplikazioaren azpiegitura ondorengo baliabideek osatzen dute: web sistemak, administrazio eta kudeaketa lanetarako moduluak, datu base kudeatzaile sistemak, disko fisikoak, galdera sortzaile sistemak, analizatzaile morfologikoak eta web zerbitzu gisa berrerabili den analizatzaile sintaktikoak; azken bi tresna hauek -analizatzaile sintaktikoa eta morfologia- IXA taldean garatutako tresnak dira. Hurrengo irudian garatutako aplikazioaren arkitektura orokorra azaltzen da, diseinuari dagokion kapituluan zehatzago aztertuko duguna.



Irudia 3.5: Aplikazioaren arkitekturaren eskema.

3.3 Erabilpen kasuen analisia

Atal honetan, aplikazioak dituen erabiltzaile mota guztiek izango dituzten beharrak analizatuko dira. Aipatu gisa, aplikazioan hiru erabiltzaile mota egon ahal dira: erabiltzaile kautotu gabeak -bisitariak-, erabiltzaile kautotuen eta sistemaren administratzailea.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

3.3.1 Erabiltzaile kautotu gabeen erabilpen kasuen analisia

Erabiltzaile kautotu gabeen kasurako, identifikatu diren beharrak ondorengoak dira:

Erabiltzaile kautotu gabeen beharrak
Webgunearen informazioa edota laguntza eskuragarri izatea
Aplikazioan erregistratzeko aukera izatea
Saioa hasteko aukera izatea
Beste erabiltzaileen testu fitxategiak irakurtzeko gai izatea
Beste erabiltzaileen testuetan idatzi diren iritziak irakurtzeko gai izatea
Beste erabiltzaileek esleitutako balorazioak ikusteko gai izatea
Eskura dauden fitxategien gainean emaitzak filtratzeko aukera izatea
Landu nahi diren galdera kopurua finkatu eta deklinabide kasuak zehazteko gai izatea
Automatikoki sortutako ariketak lantzeko gai izatea
Ariketak lantzeko tresna lagungarriak eskura izatea

Taula 3.1: Erabiltzaile kautotu gabeen beharrak identifikatzen dituen taula.

3.3.2 Erabiltzaile kautotuen erabilpen kasuen analisia

Erabiltzaile kautotuen kasurako, identifikatu diren beharrak ondorengoak dira:

Erabiltzaile kautotuen beharrak
Webgunearen informazioa edota laguntza eskuragarri izatea
Pasahitza galduz gero berreskuratzeko aukera izatea
Saioa amaitzeko aukera izatea
Profila ikusatzeko edota aldatzeko aukera izatea
Kontua ezabatzeko aukera izatea
Beste erabiltzaileekin komunikatzeko gai izatea
Testu fitxategiak aplikaziora igotzeko gai izatea
Beste erabiltzaileen testu fitxategiak irakurtzeko gai izatea
Gordetako testu fitxategiak irakurtzeko, ezabatzeko edota partekatzea aldatzeko gai izatea
Beste erabiltzaileek idatzitako iritziak ikusi eta besteen testuetan iritziak idazteko gai izatea
Beste erabiltzaileek esleituriko balorazioak ikusi eta besteen testuei balorazioak esleitzeko gai izatea
Testu fitxategi bati lotuta sortu diren galderak editatzeko gai izatea (ezabatu, aldatu, osatu, ...)
Automatikoki sortu diren galderez gain, galdera berriak eskuz gehitzeko gai izatea
Eskura dauden fitxategien gainean emaitzak filtratzeko aukera izatea
Landu nahi diren galdera kopurua finkatu eta deklinabide kasuak zehazteko gai izatea
Automatikoki sortutako ariketak lantzeko gai izatea
Ariketak lantzeko tresna lagungarriak eskura izatea
Konfiantzazko erabiltzailea bihurtzeko aukera izatea

Taula 3.2: Erabiltzaile kautotuen beharrak identifikatzen dituen taula.

3.3.3 Administrazioaren erabilpen kasuen analisia

Sistemaren administrazioaren kasurako, identifikatu diren beharrak bi multzotan banatu dira: alde batetik, erabiltzaile erregistratu batek dituen beharrak -aurreko atalean azaldutakoak-, eta bestetik, aplikazioa administrazioko beharrak. Aplikazioaren mantenurako administrazioari bakarrik dagozkion erabilpen kasuak ondorengoak dira:

Administrazioaren beharrak
Erabiltzaileak ezabatzeko gai izatea
Erabiltzaile kontuen informazio guztia ikusteko gai izatea
Erabiltzaileak dituzten fitxategi publiko eta pribatuen informazioa ikusteko gai izatea
Erabiltzaile guztien fitxategi publiko eta pribatuak irakurtzeko gai izatea
Erabiltzaileek igotako testu fitxategiak onartu eta baztertzeko gai izatea
Erabiltzaileek igotako testu fitxategiak ezabatzeko gai izatea
Erabiltzaile kautotuak konfiantzazko erabiltzaileak bilakatzeko gai izatea

Taula 3.3: Administrazioaren beharrak identifikatzen dituen taula.

Erabilpen hauek guztiak erabiltzaileak kudeatzeko txokoan bilduko dira, eta, hainbestez, txoko honetan lan egiteko administrazioa bakarrik izango da gai.

3.4 Software erabakien eta teknologiaren analisia

Atal honetan garapenean behar izango diren tresnak eta baliabideak analizatuko dira. Analisi hau burutzeak berebiziko garrantzia du aplikazioa garatzeko, erabili behar diren teknologiak bateragarriak eta elkar lanean aritzeko aproposak izan behar baitute. Honetaz gain, software garapenarekin erlazionatutako erabakiak ere azalduko dira.

3.4.1 Erabilitako azpiegitura, teknologia eta baliabideak

Ondorengo lerroetan proiektua garatzeko aukeratu diren tresnen deskribapen txikia egingo da. Tresna batzuen kasuan -Apache eta MySQL esaterakoa hauek erabiltzera behartuta egon gara IXA taldeko zerbitzarietan iada martxan daudelako, beste tresnak, ordea, hauekin leihan daudenak baino gaitasun gehiago dituztelako aukeratu dira. Hona aukeratutako azpiegitura, teknologia eta baliabideen zerrenda:



Software librearen munduan aurki daitekeen datu base sistema[5, 26] ezagunenetarikoa bat da. Errendimendu handikoa, fidagarritasun altukoa eta erabilerraza izateari atxikitzen zaio lortutako arrakasta. Hogeita hamar plataformak baino gehiagotan exekuta daiteke, hauen artean: Linux, Windows, Mac OS, Solaris, HP-UX, IBM AIX, ... Honetaz gain, MySQLko komunitateak tresna eta baliabide ugari eskaintzen ditu erabiltzaile, diseinatzaile zein garatzaileentzat.



MySQL Workbench

MySQL datu baseak diseinatzeko eta kudeatzeko baliabidea da. Taulak, errutinak, bistak, ... sortu edota editatzeko aukera ematen du. Interfaze la-

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

gungarri batez hornituta eraginkortasun handiz lan egitea ahalbidetzen du.

<?xml?> XML

XML[23] (Extensible Markup Language) markatze hizkuntza hedagarri bat da, hau da, xede orokorreko markaketa lengoaia. Beste helburu batzuetarako erabili badaiteke ere, informazio sistema ezberdinen artean egituratutako datuak partekatzeko erabiltzen da gehienbat. Izan ere, modulu ezberdinen arteko komunikazioa modu txukun eta egituratuan egitea ahalbidetzen du.



HTML 5

HTML5[4, 23] (HyperText Markup Language, version 5) World Wide Webaren oinarriko hizkuntza denaren bostgarren berrikustea da. Oraindik esperimentala izanik ere, nabigatzaile zein garatzeile askok erabiltzen dute; honek eskaintzen dituen abantaila, aurrerapen eta hobekuntzak direla eta.



CSS 3

CSS3 (Cascading Style Sheets, version 3)[23] beste teknologiek -HTML5 dagokion kasuan- egituratutako dokumentu baten aurkezpena definitzeko balioko duen lengoaia da. CSS lengoaiaren helburu nagusia dokumentu baten edukia eta aurkezpena bereiztea da.

HTML5 eta CSS3 orain arte web munduan egon den multimedia eduki falta osatzeko ahaleginak dira. Izan ere, estandar libreen multimedia sortzeko gaitasun urria zela eta, beste teknologia batzuk bilakatu ziren eduki

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

mota hauen pilareak, esate baterako: Flash teknologia. Denbora igaro ahala, oso teknologia ospetsua bilakatu zen eta segituan Flash-ez garatutako webguneak agertu ziren. Programatzaile zein diseinatzaile asko Flash teknologietan espezializatzen ziren heinean, Flash teknologiaren inguruko okerreko erabilerak ere nabarmentzen hasi ziren -guztiz Flash-ez garatutako webguneak esaterako-

Hala eta guztiz ere, badirudi azken urteetan jarrera hau baztertzen ari dela beste teknologia batzuen alde. Era honetan, estandar libreei bultzada bat ematekotan, berauek berrikusteko eta hobetzeko beharra oso nabarmena zen multimedia sortzeko gaitasunean bereziki. Honela, HTML4 eta CSS2 teknologiak berregituratu ziren[23] aipaturiko HTML5 eta CSS3 estandar berriak sortuz.

Erabilera zuzenei jarraiki, aplikazio hau estandar libreak erabiliz garatuko da aipatu den mugimendu berri hau jarraitzekotan.

JavaScript JavaScript

JavaScript[23] interpretatutako programazio lengoaia bat da, Java programazio lengoiaarekin antzekotasun handia duena, baina web orrien garapenera zuzenduta dagoena. Bezeroaren aldean -nabigatzailean- exekutatzeko den programazio hizkuntza izanik, zerbitzariaren lan karga txikitu egiten du nabarmenki. Web teknologia ezberdinen elkarlana ahalbidetzeko JavaScript erabiltzea oso garrantzitsua izango da, honen bitartez AJAX erabiltzeko aukera izango baita.



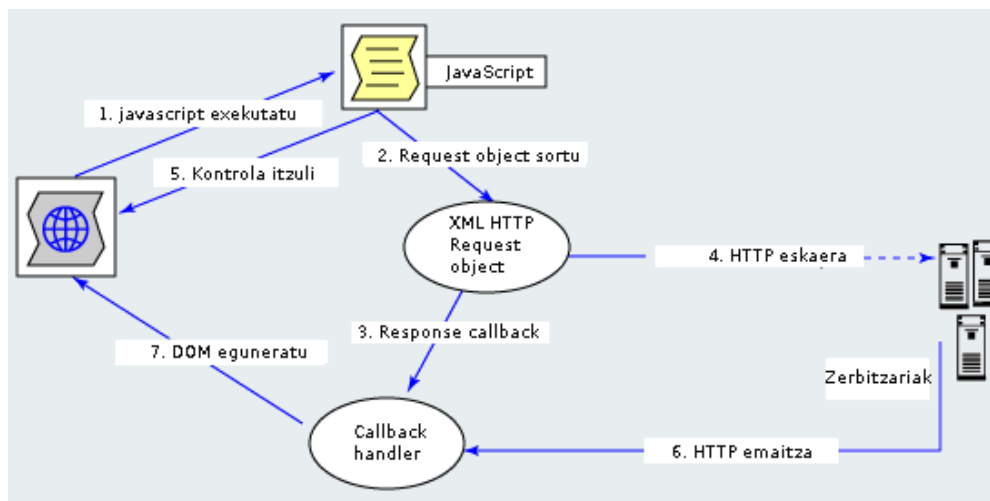
Ajax

Ajax (Asynchronous JavaScript and XML)[23, 8] JavaScript eta XML teknologiak bateratuz web interaktiboen garapenerako teknika da. Ajax-en helburua web orrialdeen abiadura, elkarreragina, funtzionalitatea eta erabilgarritasuna handitzea da.

Honetarako, Ajax-eko motoreak eskaera asinkronoak bidaltzen dizkio zerbitzariari -erabiltzailearekiko garden egiten da hau- eta berau arduratzen da web zerbitzariarekin konexioa egin eta emaitza interpretatzeaz. Era honetan,

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

web orrialdeak askoz ere harkorragoak bilakatzen dira.



Irudia 3.6: Ajax exekuzio eskema eredu.



Testu fitxategiak aztertu eta manipulatzeko bereziki diseinaturiko programazio hizkuntza da Perl[1]. Software librekoa ere bada, eta ia sistema eragile guztietan erabil daiteke. Aipagarriena adierazpen erregularrekin lan egiteko eskaintzen duen tresneriaren indarra da. Gainera, honek duen komunitate aktiboari esker modulu askorekin hornitu daitekeen hizkuntza da.



PHP

PHP[5, 25, 8, 3] (Hypertext Preprocessor) web garapenean eduki dinamikoak duten webguneak garatzeko baliagarria den hizkuntza da. Exekutatzeke errekurso gutxi behar dituen, abiadura handikoa eta egonkorra den zerbitzariaren aldeko hizkuntza da PHP. Datu base sistema ugarirekin funtzionatzeko

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

aukera izatea eta sistema eragile gehienetarako eskuragarri izatea dira beronen abantaila nagusiak.



Apache HTTP zerbitzaria[24] web zerbitzua bermatzen duen softwarea da. Oso ezaguna da software librearen munduan eta plataforma askotan erabili daitekeen arren (Unix, FreeBSD, Linux, Solaris, Novell NetWare, OS X, Microsoft Windows, OS/2, TPF, eComStation, ...), unix motako makinetan erabilia izan da gehienbat.

Apache, MySQL eta PHP hirukotea, gaur egun, webgune asko garatzeko erabiltzen da, eduki dinamikoa duten orriak sortzeko talde sendoa osatzen baitute[5].



Eclipse IDE

Eclipse IDE[29] (Eclipse Integrated Development Environment) softwarea garatzeko baliagarria den tresna bat da. Eskuragarri dituen modulu guztiei esker, programatzaileari funtzionalitate eta baliabide asko eskaintzen dizkion aplikazio bat da. Era honetan, programazio lanak nabarmen errazten direlarik.



Emacs[27] testua editatzeko aplikazio bat da. Programatzailearen ikuspuntutik ez du softwarea garatzeko IDE-k bezain besteko laguntza eskaintzen baina, script txikiak garatzeko aproposa da. Unix motako arkitekturetan oso ezaguna eta erabilia, abantaila nagusien artean hedagarria eta arina izatea dira.



Visual Paradigm (VP)[28] aplikazioen bizitza zikloko fase ezberdinak islatzeko erabiltzen den plataforma da. Aplikazio honek UML diagramak era txukun eta antolatuan sortzea ahalbidetzen du, esaterako: erabilpen kasu diagramak, klase diagramak, kolaborazio diagramak, denbora diagramak, etab.



Kile[30] TeX edota LaTeX dokumentuak garatzeko tresna baliagarria da. Aipaturiko dokumentuak garatzeko laguntza anitz eskaintzen duenez, eginbeharrak nabarmenki errazten dituen aplikazioa da. Unix edota Mac OS X motako sistematan bakarrik aurki daiteke Kile.

3.4.2 IXA taldean garatutako tresnen eta baliabideen erabilera

Eihera

Eihera[32] sistemak euskarazko entitate izendunak ezagutu eta sailkatzen dituen tresna da. Lau pausotan diseinatu da sistema: lehenengoan, ezagutzailea garatu da egoera finituzko transduktoretan errepresentatutako informazio linguistikoan oinarrituta; bigarrenean, transduktore horien emaitzetatik abiatuta erdi-automatikoki etiketatutako corpusak sortu dira; hirugarrenan, machine learning teknikak aplikatu dira corpus horien gainean ahalik eta ezagutzaile onena lortzeko, eta azkenik, lortutako ezagutzaileak konbinatu egin dira.

Eiherak hiru klasetan sailkatzen ditu ezagututako entitateak: pertsonak, erakundeak eta tokiak.

Eulia

EULIA[33] ingurunea software arkitektura bat da, lengoia naturaleko prozesatzaileak integratzeko aukera ematen duena. EULIAren ezaugarri nagusiak, besteak beste, hauek dira: galdeketak egitea, testuak analizatzea edota testuen gainean eginiko anotazioak ikusi eta haiei buruz galdetzea.

Ixati

Ixati[34] euskararako analizatzaile sintaktikoa da, entitateak eta postposizioak identifikatzeko moduluak barneratuta dituena.

Morfeus

Morfeus[35] euskararako analisi morfologikoa egiten duen aplikazioa da. Morfeusek sarrerako testua jaso eta, tokenetan banatu ondoren, horietako bakoitzarentzat lema eta morfema konbinazio posible guztiak ematen ditu; honetaz gain, hauei buruzko informazio morfologiko guztia ere ematen du.

3.4.3 Frameworken inguruko erabakia

Frameworken inguruan hartutako erabakiak azaldu aurretik, framework bat zer den definitu beharra dago.

Framework bat softwarea garatzeko tresna lagungarri bat da. Oro har, programazio lengoia batean oinarritzen dira eta honen gainean lan egiteko liburutegi multzo handi bat eskaintzen dute. Liburutegi hauen bidez softwarea garatzeko garaian laguntza handia eskaintzen diote programatzaileari, bai kudeaketa aldetik, baita garapen aldetik ere.

Ondoren, frameworkak erabiltzearen abantaila nagusiak multzokatu dira:

- Aplikazioak garatzeko denbora murriztu egiten da.
 - Software osagaiak inplementatuta daude liburutegietan.
 - Liburutegietako funtzioak eraginkorrak dira, betiere frameworkaren sendotasunaren arabera egonik.
- Software osagaien berrerabilpen handia dago.
- Arkitektura ondo definituta dago eta software osagai guztietan mantentzen da.

Hala ere, frameworken erabilerak desabantaila batzuk dakartza:

- Garatzen den aplikazioaren kodea erabat baldintzatuta geratzen da liburutegi hauekiko.
- Konputazionalki garestia da, frameworkak eskaintzen dituen software osagaiak orokorreki izaten direlako.
 - Askotan liburutegi hauetako software osagaiak ez dira guztiz bat etortzen aplikazioaren beharrekin. Hau da, behar dena baino funtzionalitate handiagoa izan ohi dute.

Aurkezten den proiektuaren helburuetako bat ikaslea gaur egungo teknologietan trebatzea izanik, frameworkak ez erabiltzea erabaki da. Era honetan, programazio lengoiaekin zuzenean lan eginez trebakuntza handiagoa lortzea espero da. Era berean, kodea ez da frameworkera baldintzatua egongo eta software osagaiak era espezifikoan diseinatu ahalko dira.

3.4.4 Txostena LaTeX bidez garatzeko erabakia

Aurkezten den karrera bukaera proiektua \LaTeX bidez idazteko erabakia hartu izanaren arrazoiak bi dira. Alde batetik, LaTeX bera ikastearen helburua, gerora begira oso erabilgarria suertatzea espero delako; eta bestetik, LaTeX lengoaiak txosten teknikoak idazteko eskaintzen dituen abantailaz baliatzeko.

LaTeX dokumentuak eratzeko sistema bat da. Gaur egun bere erabilera oso zabaldua dago tesiak, artikuluak edota liburu teknikoak idazteko, kalitate tipografiko handia eskaintzen duelako.

Abantailarik nagusien artean ondorengoak aipa daitezke: LaTeXen edukia eta edukiaren formatua banatu egiten dira. Era honetan, LaTeX dokumentu batean edukia idazten den heinean formatua definitzeko komandoak izango dira baliagarri. Honela, LaTeXek edukia lantzea baimentzen du, honek izango duen formatuaz kezkatu gabe, eta alderantziz.

Azkenik, aipatu beharra dago, karrera bukaerako proiektu hau EHUko Software Libre Taldeak (Itsas taldeak) eskaintzen duen txantilo bat jarraituta eraiki dela, honenbestez, Itsas taldeari eskertzen zaio txantilo hau partekatu izana, proiektu hau LaTeXen idaztea ahalbidetu duelako. Itsas unibertsitate mailan software librea, estandar irekiak eta kultura librearen idealak bultzatu nahi dituen EHUko Software Librearen inguruko taldea da.

4 Kapitulu

Diseinua

Gaien Aurkibidea

4.1 Galdera sortzailea	45
4.2 Lehen bertsiotik bigarren bertsiora	50
4.2.1 Sarrera corpusaren tratamendua egokitzen	50
4.2.2 Galderak eta zehar galderak baztertzen	53
4.2.3 Hautagaien aukeraketa egokitzen	53
4.2.4 Galdetzaileen esleipena egokitzen	56
4.2.5 Galderaren eraikuntza egokitzen	59
4.3 Web aplikazioaren diseinua	61
4.3.1 Deskribapena	61
4.3.2 Modulu nagusien diseinua	63
4.4 Galdera sortzaile eta web aplikazioaren arteko komunikazioa	75
4.4.1 Sarrera corpusa analizatzeko prozesua	75
4.4.2 Galderak sortzeko prozesua	76
4.5 Informazioa egituratzea eta datu basearen diseinua	80
4.6 Erabilpen kasuen diseinua	84
4.6.1 Erabilzaile kautotu gabeen erabilpen kasuen diagrama	84
4.6.2 Erabilzaile kautotuen erabilpen kasuen diagrama .	86
4.6.3 Administratzailearen erabilpen kasuen diagrama .	89
4.6.4 Erabilpen kasuen fluxu diagramak	93
4.7 Segurtasuna bermatzen	102

4.7.1	Fitxategi sistema babesten	103
4.7.2	Pasahitzak babesten	104
4.7.3	Erabiltzaileen fitxategi pribatuak babesten	105
4.7.4	SQL injection-en aurka babesten	108

Kapitulu honetan diseinuari dagozkion atalak aurkeztuko dira; analisisian azaldutako kontzeptuak nola gauzatu diren azalduz. Lehendabizi, proiektuaren abiapuntu den galdera sortzailearen diseinua azalduko da, eta, halaber, lehen bertsio honen bilakaera. Ondoren, web aplikazioaren diseinua ikusgai jarriko da. Web aplikazioa diseinatzeko garaian ondokoak azalduko dira: datu base eta klase diagramaren diseinua, erabilpen kasuen diseinua, aplikazioaren diseinua eta hainbat segurtasun konturen erabakiak. Honetaz gain, proiektuaren atalik garrantzitsuenen arteko informazio trukearen diseinua ere azalduko da, hots, galdera sortzaile eta web aplikazioaren arteko komunikazioa.

4.1 Galdera sortzailea

Atal honetan, galdera sortzailearen deskribapen zehatza egingo da, eta, honetaz gain, lan poltsan garatutako galdera sortzailearen lehen bertsioari egin beharreko aldaketak ere azalduko dira. Proiektuaren helburu dokumentatua aipatu den modura, abiapuntua lan poltsan ikasleak inplementatutako aplikazioa da. Aplikazioa hau esaldi mailan galderak masiboki sortzeko gai den galdera sortzaile automatiko baten lehen bertsioa da.

Galderen sorkuntza

Itziar Aldabek zenbakien inguruan emandako galdera sorkuntza automatikoaren inguruko urratsei jarraiki[13], eta bere tesian[14] aurkezten den arkitektura oinarritzat hartuz, lau urratseko *Question Generation* (QG) sistema inplementatu zuen ikasleak proiektua hasi aurretik egindako lan poltsan zehar. Sistemak inplementatzen dituen lau urratsak ondorengoak dira: sarrera fitxategia prozesatzea, hautagaiak aukeratzea -galderen emaitza izango den edukiaren aukeratzea alegia-, galdetzaileak identifikatzea eta galdera eraikitzea.

Gaia identifikatu aurretik -hautagaiak aukeratzeko ataza-, sarrera fitxategi gisa erabiltzen den corpusa morfosintaktikoki analizatzen da. Era honetan, Ixatik[34] esaldien informazio linguistikoaz hornitzen du galdera sortzailea. Informazio linguistiko hau guztia *Kyoto Annotation Framework* (KAF) [11] bidez adierazten da.

Garatutako sistemaren helburuetako bat esaldi ez anbiguoetan azaltzen diren gaien buruz galderak sortzea izanik, hautatzen diren esaldiak aditz jokatu bakarria izan behar dute. Aditz jokatu eta jokatu gabeak identifikatzeko erabiltzen den eskema hau da:

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Part of Speech patroia	Aditz mota
V.ADI-(.*) + V.ADI-ADL	Aditz jokatua
V.ADT	Aditz jokatua
V.ADI-ADK	Aditz jokatua
V.ADI-SIN	Aditz jokatu gabea

Taula 4.1: Aditz jokatu eta jokatu gabeak identifikatzeko eskema.

Sarrera testua prozesatu

Lehen fase honetan galdera sortzaile sistemarentzat beharrezko diren ezau-garri linguistikoak biltzen dira. Oro har, prozesu hau atal desberdinetan banatzen da. Lehenbizi, hizkuntza naturalaren prozesamendurako (NLP) tresnak[32][33][34][35] erabiliz sarrera corpora linguistikoki analizatzen da. Garatu den galdera sortzaile sistemaren kasuan *txt2kaf.eu* Perl aplikazioa erabiltzen da. Tresna hau IXA taldeak garatu duen analizatzaile bat da, eta NLP tresnen katea aplikatzen dio adierazitako sarrera corpusari. Tresnaren izenak dion modura, sarrera corpora testu fitxategi gordina da, eta, aldiz, emaitza *Kyoto Annotation Framework* (KAF) [11] erabiliz itzultzen da.

Behin sarrera testua analizatu ondoren, galdera sortzaile sistemak informazio linguistikoa biltzen duen emaitza fitxategia parseatzen du. Era honetan, NLP tresnetatik abiatuta lortu dugun informazio guztia jasotzen du sistemak, eta hurrengo atazekin hasteko gai da. Irakurketa hau guztia XML::LibXML Perl moduluaren laguntzaz[7] egiten da.

Hautagaiak aukeratu

Bigarren fasea, galderen emaitzak izango diren osagaiak hautatzeaz arduratzen dena. Ataza hau burutzeko prozesatua izan den sarrera testua esaldiz esaldi aztertzen da eta KAF formatuak eskaintzen dituen baliabideak erabiliz galderen emaitza izango diren sintagmak identifikatzen dira.

Galderen emaitzak izango diren hautagaiak identifikatzeko irizpidea lehiaketa batean oinarritzen da; lehiaketa honen kandidatuak esaldiko izen sintagma guztiak izanik. Era honetan, esaldi bakoitzerako, esaldi horretako sintagma guztiak sekuentzialki prozesatzen dira, eta deklinabide kasu bakoitzerako aukeratzen den hautagaia informazio linguistikoan oinarrituta pisu handiena lortu duena izango da. Ondorioz, aurkitzen den deklinabide kasu bakoitzeko galdera bat sortuko da. Heuristiko¹ honen atzean dagoen ideia ondorengoa da: informazio linguistikoan oinarritutako pisutze sistema bat

¹Arazo jakin bat ebazteko diseinatzen den teknika da heuristikoa

ezartzea. Ondorioz, aurretik aurreratu den moduan, garatutako galdera sortzailea esaldi mailan galderak masiboki sortzeko gai den sistema da.

Galderak sortzeko sistema honetan, pisuak esleitzeko informazio iturri desberdinak erabiltzen dira: alde batetik, izen sintagmen informazio morfolo-
gikoa, eta, bestetik, NLP tresnak identifikatzeko gai diren entitateak -tokiak,
pertsanak eta lekuak-. Informazio linguistiko hau guztia bide desberdinetatik
abiatuta lortzen da:

- Entitateak

Ixatik [34] entitateak hautemateko eta klasifikatzeko tresna bat in-
tegratzen du: Eihera [32]. Garatutako sistemak tresna hau erabiltzen
du pertsonak, tokiak eta lekuak identifikatzeko.

- Morfologia

Izen sintagmen deklinabide kasuari buruzko informazioa analiza-
tzaile sintaktikoak[34] erauzten du. Sistema identifikatzeko gai den
deklinabide kasuak ondorengoak dira: ablatiboa, adlatibo bukatuzkoa,
adlatibo bide zuzeneko, adlatiboa, soziatiboa, datiboa, destinatiboa,
ergatiboa, genitibo leku denborazkoa, genitiboa, inesiboa, instrumen-
tala, motibatiboa, absolutiboa, partitiboa, prolatiboa, banakaria eta
deskribatzailea.

Ondorioz, metodologia honek automatikoki erauzitako informazioaz ba-
liatzen da hautagaiei informazio linguistiko atxikitzeke.

Hautagaien aukeraketari dagokionez, sekuentzialki exekututzen den heu-
ristikoan datza. Aurretik aipatu den modura, heuristikoaren hautagaiak
esaldiko izen sintagma guztiek osatzen dute. Heuristiko honi jarraituz es-
kuratu diren balizko hautagai guztien artean deklinabide kasu bakoitzeko
bakarra hautatzen da. Heuristikoaren funtzionamendua txapelketa modukoa
da. Hau da, heuristikoak hautagai posibleen arteko alderaketa bat egiten du
informazio linguistikoan oinarrituta eta azkenean pisu handiena duen hauta-
gaia aukeratzen du. Hurbilpen honetan, heuristikoak entitateei lehentasuna
ematen die:

1. Izen sintagma entitate bat baldin bada -pertsone, tokia edo lekua- sis-
temak bi puntutako kalifikazioa ematen dio.
2. Morfologia, kasu guztietan, puntu batekin kalifikatzen da.

Izan bedi A lehen hautagaia, B bigarren hautagaia eta Pisu(X) X hauta-
gaiari esleitutako pisua. Erabiltzen den heuristikoa ondoko taulan laburbil-
tzen da:

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen
aplikazioa

Egoera	Heuristikoak aukeratzen duen hautagaia
$Pisu(A) > Pisu(B)$	A hautagaia
$Pisu(A) < Pisu(B)$	B hautagaia

Taula 4.2: Hautagaien aukeraketa heuristikoa. Lehen bertsioan

Galdetzaileak aukeratu

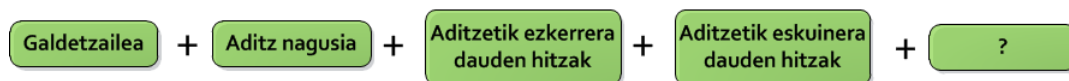
Hautagaiak identifikatu ostean, prozesuaren hirugarren atalari ekiten dio aplikazioak. Ataza honen helburua galdetzailearen edo galdera motaren identifikazio zuzena da. Galdetzaileen aukeraketa hautatu diren izen sintagmen informazio linguistikoaren arabera egiten da. Honela, ezaugarri linguistiko bakoitzeko -entitateak eta morfologia-, hizkuntzalariek ezarritako galdetzaileak esleitzen dira, dagozkien izen sintagmentzako probabilitate handiz zuzenak izateko estimatu direnak.

Galdetzaileak aukeratzeko fase honetan, hurrengo deklinabide kasuak bakarrik kontuan hartzen dira: ablatiboa, adlatibo bukatuzkoa, adlatibo bide zuzeneko, adlatiboa, soziatiboa, datiboa, ergatiboa, genitibo leku denborazkoa, genitibo, inesiboa eta absolutiboa. Izan ere, galderak sortzeko sistema trebatzeko erabili den corpusean kasu hauek dira agerpen kopuru handiena dutenak.

Galdera eraiki

Azkenik, aplikazioak galderaren sorkuntzari ekiten dio. Fase hau inplementatzeko algoritmo bat erabiltzen da, esaldiko sintagmak berrantolatzen eta esaldiari galderazko zentzua ematen diona; ahalik eta galdera naturalena lortzeko helburuarekin. Honetarako, hurbilpen honetan[13] erabiltzen den antzeko ebazpen bat inplementatu da.

Ondorioz, galdera eraiki aurretik jatorrizko esaldian prozesatze lana egin behar da, era honetan, lehenbizi esaldiko puntuazio ikurrak ezabatzen dira, eta, ondoren, galdera eraikitzeke patroia martxan jartzen da. Eraldaketa erregela honen errepresentazio grafikoa ondorengo irudian azaltzen da:

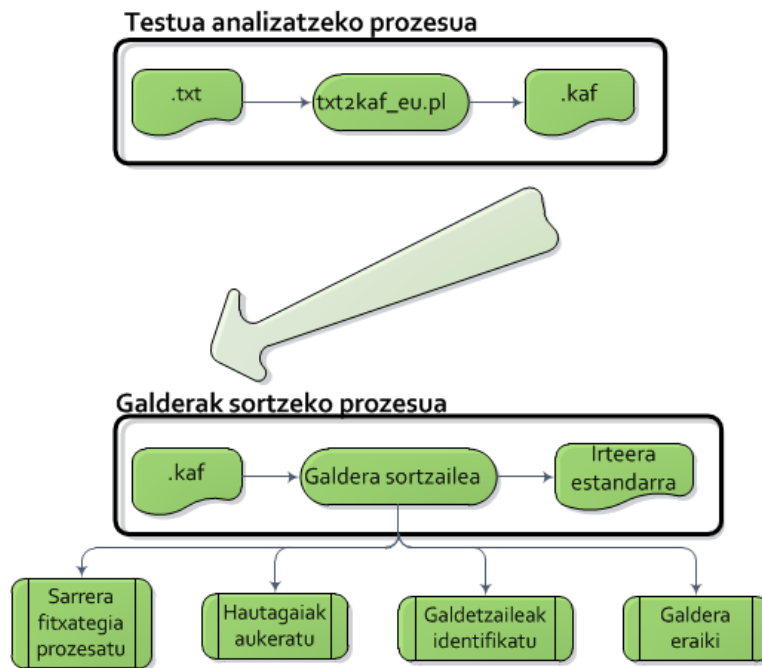


Irudia 4.1: Eraldaketa erregelaren errepresentazio grafikoa.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Laburpena

Galderak automatikoki sortzeko sistemaren lehen bertsioa deskribatzen duen eskema logikoa ondorengo irudian ikus daiteke.



Irudia 4.2: Galdera sortzailearen hasierako eskema logikoa.

4.2 Lehen bertsiotik bigarren bertsiora

Lan poltsan inplementatutako galdera sortzaile sistemak hainbat aldaketa eskatzen zituela kontuan hartuz, lehen bertsioaren gainean lan egin da bigarren bertsio bat lortzeko. Bigarren bertsio hau lehen bertsioaren bilakaera bat da, zeinetan hainbat hurbilpen linguistikotan oinarritu garen -diseinatu behar izan diren hainbat egokitzapenez gain- galderak sortzeko sistema hobetzeko asmoz. Atal honetan, galdera sortzailearen lehen bertsioak izandako bilakaera deskribatuko da.

4.2.1 Sarrera corpusaren tratamendua egokitzen

Aurretik aipatu den modura, lehen fase honen helburua analizatzaile linguistikoez baliatzea da galdera sortzaile sistema informazio linguistikoz hornitzeko. Hortaz, fase honetan bi ataza nagusi bereiz daitezke: bata, analizatzaile linguistikoak erabiltzea, eta bestea, informazioa parseatzea. Galdera sortzaile sistemaren bigarren bertsioan egindako aldaketak bi ataza hauen optimizazioa eta egokitzapena izan dute helburu.

Sarrera corpusaren prozesaketa

Galdera sortzailearen (QG) lehen bertsioan gertatzen ez zen modura, proiektu honetan QG sistema web aplikazio baten motor gisa erabiliko da galderak automatikoki sortzeko. Hau dela eta, lehen bertsioan diseinatutako sarrera ez da baliagarria izango, izan ere, komunikazio eskema berrian web aplikazioak bere kabuz sarrera corpusa linguistikoki analizatu beharko duelako, eta, ondoren, fitxategi hau QG sistemari eman. Behar berri hauek direla eta, ez da posible izango orain arte corpusak analizatzeko erabili den aplikazioa berrerabiltzea, eta, honen ordez, analisi linguistikoa egiteko gai den web zerbitzu bat erabiliko da. Dena den, aldaketa hauek *Galdera sortzaile eta web aplikazioaren arteko komunikazioa* azpiatalean (4.4) sakonago azaltzen dira, galderen sorkuntzarekin baino web aplikazio eta QG sistemaren arteko komunikazioarekin lotuta baitaude.

Analizatzaile sintaktikoaren hutsegiteak leuntzen

Sarrera corpusaren informazio linguistikoa lortzeko egiten den analisia automatikoki egikaritzen denez, noizean behin inkonsistentziaren² bat edo beste sortzen da. Akats hauek galderen sorkuntzan eragin zuzena dutenez, hauek eragindako kaltea mugatzeko asmoarekin inplementatu da QG sistemaren

²Eranskinetan *Analisi linguistikoan arazoak* atalean (I.a) arazo hauen berri ematen da.

bigarren bertsioa. Hala ere, inplementatu diren aldaketak azaldu aurretik, ezinbestekoa da *Kyoto Annotation Framework* (KAF) [11] zehaztapenean sakontzea.

*KAF*en helburua anotazio sintaktikoak eta semantikoak adierazteko baliagarria den oinarriko formatu bat definitzea da. Proiektuan azaldu den modura, QG sistemak erabiltzen dituen analizatzaile sintaktikoen emaitzak formatu honetan adierazten dira, eta formatu honen elementuak dira hain zuzen, egindako aldaketak behar bezala ulertzeko azaldu behar direnak. Hona hemen KAF formatuan aurki daitezkeen elementurik garrantzitsuenak:

- Word form nodoak

Tokenizazioaren ondoren lortzen diren elementuak.

Zein esaldi identifikadoreetatik datozen adierazten da hauen atributuen artean.

Ondorengoak word form nodoen adibideak dira:

```
<wf wid="w2" sent="1">ikasle </wf>
<wf wid="w3" sent="1">gazte </wf>
<wf wid="w4" sent="1">batek </wf>
```

- Term nodoak

Termino nodoek aurretik azaldutako word form nodoetara -edo multi word form nodoetara- erreferentziak egiten dituzte.

Termino nodoek, besteak beste, ondorengo atributuak biltzen dituzte: lemma, part of speech (pos) eta name entity (netype).

Ondorengoak termino nodoen adibideak dira:

```
<!-- ikasle -->
<term tid="t2" type="open" lemma="ikasle" pos="N.IZE-ARR">
<span>
<target id="w2"/>
</span>
</term>
<!-- goitik behera -->
<term tid="t80" type="open" lemma="goitik_behera" pos="A.ADB-ARR">
">
<span>
<target id="w91"/>
<target id="w92"/>
</span>
</term>
<!-- Parisera -->
```

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

```
<term tid="t97" type="entity" lemma="Paris" pos="R.IZE-LIB" case
  ="ALA" netype="Tokia">
<span>
<target id="w113"/>
</span>
</term>
```

- Chunk nodoak

Chunkak termino nodoetara erreferentziak biltzen dituzten sintagmak dira.

Hauen atributuen artean ondorengoak aurki daitezke: head, phrase eta case.

Ondorengoak chunk nodoen adibideak dira:

```
<!-- Tolosako ikasle gazte batek -->
<chunk cid="c1" head="t2" phrase="NP" case="ERG">
<span>
<target id="t1"/>
<target id="t2"/>
<target id="t3"/>
<target id="t4"/>
</span>
</chunk>
<!-- gurasoei -->
<chunk cid="c2" head="t5" phrase="NP" case="DAT">
<span>
<target id="t5"/>
</span>
</chunk>
```

Edozein kasutan, wordform nodoak, termino nodoak eta chunk nodoak identifikatzaileen bidez adierazten dira -unibokoki-; eta analisi linguistikoa erreferentzia hauek mantentzen dira. Era honetan, KAF fitxategiaren elementu guztiak erlazonaturik daude analisi linguistikoaren emaitzetan.

Hasieran planteatutako ideari helduz, analisi sintaktikoa hurrengo arazoak hauteman dira:

- Esaldi, wordform, termino eta chunk erreferentziak ez datoz bat jatorrizko corpusetako testu gordinarekin.
 1. Tokenizazio prozesua amaitzean hitz bat bi aldiz -edo gehiago- analizatu da.
 2. Tokenizazio prozesua amaitzean hitz bat edo beste ez da analisi linguistikoa ageri.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Galdera sortzaile sistemaren bigarren bertsioan inplementatu den kontrol egiturari esker, bi egoera hauek detektatzen eta tratatzen dira. Izan ere, lehen kasuari dagokionez, tokenizazio prozesua amaitzean analisisian hitz bat bi aldiz agertzen bada, esaldia baztertu egiten da. Bigarren kasuari dagokionez, tokenizazio prozesua amaitzean hitz bat analisi linguistikoan ageri ez bada, hitz horri erreferentzia egiten dion alegiazko elementua sortzen da -chunk birtual izena eman dioguna- eta, era honetan, esaldia ez da baztertu eta galderen sorkuntza prozesuarekin jarraitzeko moduan da.

Ondoren alegiazko elementu hauek sortzeko erabiltzen den eskemaren laburbildua ageri da:

Termino	Chunk birtuala	Termino	Chunk birtuala
Part of Speech	Phrase	Case	Case
V.AD(.*)	VP	(.*)	Zuzenean kopiatu
V.AD(.*) ez dena	NP	(.*)	Zuzenean kopiatu

Taula 4.3: Alegiazko elementuak sortzeko eskema.

4.2.2 Galderak eta zehar galderak baztertzen

Oro har, ez da ideia ona galdera -galdera zuzena, '?' ikurra duena- edo zehar galdera bati buruzko galdera berri bat sortzea. Horregatik, jatorrizko esaldia galdera edo zehar galdera gisa detektatzen bada, zuzenean baztertu egiten da. Aipatu beharra dago, KAF fitxategietan eskuragarri dugun informazioa mugatua dela, eta, horregatik, zehar galdera guztiak behar bezala identifikatzea -aditzen analisisian oinarrituta- ezinezkoa dela. Horregatik, momentuz, zehar galderetan oso maiz agertzen diren hitzak bilatuz identifikatzen ditugu zehar galderak.

Oro har, sarrera corpusaren tratamenduan diseinatu eta inplementatu diren egokitzapenak sarrera corpusaren estaldura hobe egitea dute helburu, honela, QG sistemak ziurtasun handiagoz ekiten dio hurrengo faseari.

4.2.3 Hautagaien aukeraketa egokitzen

Galderen sorkuntza prozesuaren bigarren atalean bi hobekuntza gehitu zaizkio QG sistemaren lehen bertsioari: alde batetik, bizidun eta bizigabe ezauzgarriak biltzen dituen hiztegia, eta bestetik, roll semantikoak. Hala eta guztiz ere, QG sistemaren heuristikoen funtzionamendua berdina izango da

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

bigarren bertsio honetan, hots, lehiaketan oinarritutakoa. Dena den, hurbilpen linguistiko berrien erabilerak heuristikoa inkrementalki egokitzera behartu gaitu. Ondorengo lerroetan aipatutako hurbilpen linguistikoak sakonago azalduko dira.

Bizidun eta bizigabeen hiztegia

QG sistema [17]n egindako erauzpen lan erdiautomatikoan oinarritzen da ohiko izenen informazio semantikoak eskuratzeko. Besteak beste, biziduntasun ezaugarriak dituen hiztegi hau IXA taldeko itzulpengintza sailean erabiltzen da. Hiztegi honetan hainbat semantika ezaugarri aurki daitezke analizatu diren elementuentzat, esaterako: biziduntasuna, hizkuntza, denbora, gaia edo materiala, gailua edo ibilgailua eta komunikazio tresneria. Hiztegi honen laguntzarekin³ QG sistemak 15000 izenen ezaugarri semantikoak kontuan hartzen ditu.

Hona hemen hiztegi honen hainbat adibide:

...
aztergai [BIZ-]
aztergailu [BIZ-]
azterketa [BIZ-]
azterlan [BIZ-]
azterlari [BIZ+]
azterpen [BIZ-]
azterraldi [BIZ-]
aztertzaile [BIZ+]
...
dantza [BIZ-]
dantzaldi [BIZ-]
dantzazale [BIZ+]
dantzaleku [BIZ-]
dantzatoki [BIZ-]
...

Roll semantikoak

QG sistemak predikatu mailan, diskurtso egituran eta roll semantikoetan eskuz etiketatutako corpus[20] batean oinarrituta, aditzen eta izen sintagmen arteko roll dependentziak estimatzeko gai da. Corpus hau [21]n egindako lanean oinarritzen da, euskarazko 100 aditzen inguruko analisi sakona biltzen duena. 100 aditz hauen agerpen kopuruan oinarrituta VerbNet[22]eko hurrengo rollekin lan egin da: actor, attribute, agent, beneficiary, cause, desti-

³hiztegia /sc01a4/users/ilopez077/proiekttoa/semantika_hiztegia katalogoan aurkitzen da eskuragarri.

nation, direction, experiencer, extent, instrument, location, manner, patient, predicate, product, recipient, source, theme, temporal eta topic.

Ondorengo lerroetan aipatu den baliabidearen inguruko adibide batzuk azaltzen dira; esaterako, *abestu* aditzari dagokionez deklinabide kasua inesi-boa den kasuetan roll semantikoa location motakoa da %66 batean -aurrerago aipatuko bada ere, %75eko balioa konfiantzazko muga gisa hartu da-, gainera, deklinabide kasu eta roll zehatz honekin bi kasu agertu dira ikasketarako erabili den corpusean. Hona hemen adibide gehiago:

```

...
patterns.abestu-sing_01.txt
|erg||agent|          2          100%  (|abestu:sing_01|)
|ine||location|      2           66%   (|abestu:sing_01|)
|ins||manner|        1          100%  (|abestu:sing_01|)

patterns.afaldu-have_dinner_01.txt
|emen||temporal|     1          100%  (|afaldu:have_dinner_01
|)
|ine||temporal|      1          100%  (|afaldu:have_dinner_01
|)

...

patterns.jokatu-act_01.txt
|abs||predicate|     1          100%  (|jokatu:act_01|)
|aurk||predicate|   2          100%  (|jokatu:act_01|)
|emen||predicate|   1          100%  (|jokatu:act_01|)
|erg||predicate|    2           66%   (|jokatu:act_01|)
|ins||predicate|    5          100%  (|jokatu:act_01|)
|kaus||temporal|    1          100%  (|jokatu:act_01|)
|mod||manner|       1          100%  (|jokatu:act_01|)

...

```

Hautagaiak aukeratzeko heuristikoa

Galderak sortzeko sistemaren bigarren bertsioan hurbilpen linguistiko berriak gehitu ondoren, informazio linguistikoan oinarrituta ezartzen den pisutze sistema berria hau da. Kontuan izan behar da, lehen bertsioan gertatzen zen modura, bertsio honetan ere entitateak eta biziduntasuna lehenesten dela egindako ebaluaketan beste hurbilpen linguistikoak baino emaitza hobek izan dituelako. Hona hemen pisutze sistema:

1. Izen sintagma entitate bat baldin bada -pertsonek, tokia edo lekua- sistemak hiru puntutako kalifikazioa ematen dio.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

2. Izen sintagma bizidun edo bizigabe hiztegian aurkitzen bada, kasu honetan ere, sistemak hiru puntutako kalifikazioa ematen dio.
3. Izen sintagma eta aditzaren arteko rola identifikatzea lortzen bada, sistemak bi puntutako kalifikazioa ematen dio.
4. Izen sintagmaren informazio morfologiko hutsak, edozein kasutan, puntu bateko kalifikazioa lortzen du.

Hurbilpen linguistiko berri hauei esker, QG sistemaren hautagaien aukeraketa prozesuak bi aurrerapen nagusi izango ditu: bata, hautagaiak zehaztasun handiagoz identifikatzeko gai izango dela, eta bestea, identifikatzen diren hautagaien inguruan informazio linguistiko handiagoa izanik, hauen arteko aukeraketa egokiago bat egiteko gai izango dela.

Gainera, pisu bereko hautagai ezberdinen arteko aukeratze liskarrak konpontzeko, hautagai eta aditzaren arteko distantzia neurtzen da, eta distantzia txikienera legokeen hautagaiak pisu handiagoa lortuko luke beste hautagaia-aren aurrean. Izan bedi A lehen hautagaia, B bigarren hautagaia eta $Pisu(X)$ X hautagaiari esleitutako pisua. Heuristikoaren bertsio berritua ondoko taulan laburbiltzen da:

Egoera	Heuristikoak aukeratzeko duen hautagaia
$Pisu(A) = Pisu(B)$	Aditzetik gertuen dagoena
$Pisu(A) > Pisu(B)$	A hautagaia
$Pisu(A) < Pisu(B)$	B hautagaia

Taula 4.4: Hautagaien aukeraketa heuristikoa. Bigarren bertsioan

4.2.4 Galdetzaileen esleipena egokitzen

Hautagaiak identifikatu ondoren hasten den fase honetan, galdetzailearen edo galdera motaren identifikazio zuzenerako egokitzapenak kontuan hartu ditugu. Fase honetan ondorengoak egin dira: galdetzaileen orrazketa orokorra eta aurreko atalean aipaturiko roll semantikoaren baliabidea galdetzailez hornitzea.

Galdetzaileen orrazketa orokorra

Galdetzaileen esleipenari dagokionez, zehaztasun handiagoa izateko lehen bertsioan QG sistemak esleitzen zituen galdetzaileak berrikusi dira. Berrikuspen hau IXA taldeko hizkuntzalariek egin dute, eta, era honetan, esleitu

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

beharreko galdetzaileak orraztu dira. Hona hemen galdera sortzaile sistemen bigarren bertsioa esleitzeko gai den galdetzaileen zerrenda.

Sintagmaren deklinabide kasua	Biziduna	Bizigabea
Ablatiboa	Norengandik	Nondik
Adlatibo bukatuzkoa	Norenganaino	Noraino
Adlatibo bide zuzenekoa	Norantz	Norantz
Adlatiboa	Norengana	Nora
Soziatiboa	Norekin	Zerekin
Datiboa	Nori	Zeri
Destinatiboa	Norentzat	Zerentzat
Ergatiboa	Nork	Zerk
Genitibo leku denborazkoa	Noren	Nongo
Genitiboa	Noren	Zeren
Inesiboa	Norengan	Non/Noiz
Absolutiboa	Nor	Zer

Taula 4.5: Galdetzaileen zerrenda.

Roll semantikoaren baliabidea

Roll semantikoek aditz sorta baten inguruko joera eta erabilera estatistikak biltzen dituztenez, informazio honetaz gain, aditz eta deklinabide kasu bakoitzeko erabili beharreko galdetzaileak gehitu dira, betiere, gutxienez estimatutako roll semantikoa %75eko ziurtasunez eman bada, balio hau konfiantzazko eremu gisa hartu baitugu. Era honetan, QG sistemak roll semantiko bat identifikatzen duenean, informazio berri hau kontsultatuz esleitu beharreko galdetzailea aurkitzen du era automatikoan. Baliabide linguistiko berri hau hurrengo helbidean atzitu daiteke⁴.

Ondorengo lerroetan fitxategi honen hainbat sarrera aurkitzen dira adibide gisa, esaterako, *abestu* aditzaren kasuan erauzi den informazioa ergatiboari bakarrik dagokiona izan da, deklinabide kasu instrumentala eta inesiboa baztertuak izan direlako, izan ere, deklinabide kasu instrumentala ez da QG sistema honetan lantzen, eta, aldiz, deklinabide kasu inesiboari atxikitzen zaion roll semantikoa ez da konfiantzazko eremura iristen⁵.

⁴/sc01a4/users/ilopez077/proiekttoa/roll_semantikoak/aztertutako_aditzak.txt

⁵4.2.3 atalean aipatutako modura: —ine—location— 2 66% (—abestu:sing.01—)

Hona hemen hainbat adibide gehiago:

```

abestu  ERG=agent=Nork
afaldu  INE=TMP=Noiz
agertu  ABS=Theme=Zer  INE=LOC=Non  SOZ=MNR=Zerekin  ERG=agent=Nork
ahantzi ABS=Topic=Zer
ahaztu  ABS=Theme=Zer  DAT=agent=Nori  GEN=Theme=Zer  SOZ=Theme=
      Zerekin
aldatu  ABS=Theme=Zer  ERG=agent=Nork  INE=TMP=Noiz
amaitu  ABS=Theme=Zer  ABL=MNR=Nola  ALA=MNR=Nola  ERG=agent=Nork
...
bazkaldu ERG=agent=Nork
baztertu ABS=Theme=Nor/Zer
besarkatu ABS=Theme=Nor/Zer
...
deitu  ABL=LOC=Nondik  ALA=LOC=Nora  DAT=Beneficiary=Nori  ERG=
      agent=Nork  ABS=Predicate=Zer  DAT=Theme=Nori/Zeri  INE=LOC=Non
edan  ABL=Source=Nondik  ABS=Patient=Zer  ERG=agent=Nork  INE=LOC=
      Non
eduki  ABS=Theme=Zer
egin  ABS=product=Zer  ALA=DIR=Nora  DAT=Beneficiary=Nori  ERG=agent
      =Nork  SOZ=Material=Zerekin  GEN=MNR=Nola
egokitu ABS=Theme=Zer  ALA=destination=Nora/Zertara  SOZ=MNR=
      Zerekin
...

```

4.2.5 Galderaren eraikuntza egokitzen

Galdera sortzaile (QG) sistemak inplementatzen duen galderak eraikitzeko algoritmoa aldatu ez bada ere, galderak itzultzeko erabiltzen den formatua aldatu egin da bigarren bertsioan. Izan ere, QG sistemak emaitzak sortu ondoren web aplikazioari itzuli beharko dizkio honek biltegira ditzan. Komunikazio eskema berri hau dela eta, lehen bertsioan diseinatutako irteera formatua ez da egokia behar hauei erantzuteko, eta, honen ordez, XML markatze lengoia bidez definitutako zehaztapena erabiliko da informazioa trukatzeko. *Sarrera corpusaren prozesaketa* azpiatalean (4.2.1) aipatu den modura, aldaketa hauek *Galdera sortzaile eta web aplikazioaren arteko komunikazioa* azpiatalean (4.4) sakonago analizatuko dira, galderen sorkuntzarekin baino web aplikazio eta QG sistemaren arteko komunikazioarekin loturik baitaude.

Adibide gisa ondorengo lerroetan hainbat galdera biltzen dituen XML fitxategia azaltzen da:

```
<?xml version="1.0" encoding="utf-8"?>
<galderak jatorri_fitxategia="nire_aitona_Balentin_ipuina_LH1.kaf">

  <galdera kasua="Adlatiboa" heuristikoa="Roll semantikoa(
    jaitsi_ALA_LOC_Nora)" jatorrizko_esaldi_id="3">
    <sortutako_galdera>Nora jaitsi zen aurreko asteazkenean?</
      sortutako_galdera>
    <jatorrizko_esaldia> aurreko asteazkenean herrira jaitsi
      zen</jatorrizko_esaldia>
    <galderaren_emaitza garrantzia="2"> herrira </
      galderaren_emaitza>
    <aditz_nagusia chunka="c7"
      hautagaiarekiko_distantzia_chunketan="1"> jaitsi zen</
      aditz_nagusia>
  </galdera>

  <galdera kasua="Adlatiboa" heuristikoa="Biziduna"
    jatorrizko_esaldi_id="4">
    <sortutako_galdera>Norengana eman zion amonak dirua
      joateko eta hagina kentzeko?</sortutako_galdera>
    <jatorrizko_esaldia> amonak dirua eman zion
      dentistarengana joateko eta hagina kentzeko</
      jatorrizko_esaldia>
    <galderaren_emaitza garrantzia="3"> dentistarengana </
      galderaren_emaitza>
    <aditz_nagusia chunka="c10"
      hautagaiarekiko_distantzia_chunketan="1"> eman zion</
      aditz_nagusia>
  </galdera>
```

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

```

<galdera kasua="Inesiboa" heuristikoa="Roll semantikoa(
  jaitsi_INE_TMP_Noiz)" jatorrizko_esaldi_id="3">
  <sortutako_galdera>Noiz jaitsi zen herrira?</
    sortutako_galdera>
  <jatorrizko_esaldia> aurreko asteazkenean herrira jaitsi
    zen</jatorrizko_esaldia>
  <galderaren_emaiza garrantzia="2"> aurreko asteazkenean
    </galderaren_emaiza>
  <aditz_nagusia chunka="c7"
    hautagaiarekiko_distantzia_chunketan="2"> jaitsi zen</
    aditz_nagusia>
</galdera>

<galdera kasua="Absolutiboa" heuristikoa="Morfologikoa"
  jatorrizko_esaldi_id="39">
  <sortutako_galdera>Nor/Zer ezin dut?</sortutako_galdera>
  <jatorrizko_esaldia> ezin dut gehiago</jatorrizko_esaldia>
  <galderaren_emaiza garrantzia="1"> gehiago </
    galderaren_emaiza>
  <aditz_nagusia chunka="c136"
    hautagaiarekiko_distantzia_chunketan="1"> ezin dut</
    aditz_nagusia>
  </galdera>
</galderak>

```


4.3 Web aplikazioaren diseinua

Atal honetan web aplikazioaren deskribapen osoa eta berau osatzen duten moduluak aurkeztuko dira. Hala eta guztiz ere, deskribapen hauetan aipatzen diren erabilpen kasuen fluxu diagramak hurrengo ataletan sakonago aztertuko dira.

4.3.1 Deskribapena

Web aplikazioa erabiltzen duten bezeroen helburua ariketak egitea izanik ere, bi erabiltzaile mota bereizten dira tresna erabiltzeko garaian. Alde batetik, erabiltzaile kautotuak, hau da, sisteman erregistratu diren erabiltzaileak; eta bestetik, bisitariak, azken hauek altan eman ez direnak. Erabiltzaile mota guztiek erabili ahalko dute web aplikazioa testuak lantzeko nahi adina ariketa landuz. Funtzionalitateak, ordea, mugatu egingo dira bisitarien kasuan, izan ere, gonbidatuak edota noizbehinkako bisitariak dira, eta sistemarekiko elkarrekin maila baxuagoa izango dute. Era honetan, bisitariak ezin izango dute idatzizko testurik gorde aplikazioan eta sisteman atzigarri dauden fitxategi publikoak soilik erabili ahal izango dituzte. Erabiltzaile erregistratuak, aldiz, beraien testuekin hornitu dezakete aplikazioa nahi adina idazlan igoz. Halaber, igotako testuei lotuta automatikoki sortu diren ariketak kudeatzeko aukera izango dute, esaterako: ariketak ezabatu eta editatu edo ariketa berriak sortu. Gainera, beraiek sortutako materialaz gain, material publikoa ere erabiltzeko aukera izango dute.

Honetaz gain, erabiltzaileak zein fitxategiak kudeatzeko rol berri bat agertzen da: sistemaren administratzailea. Sistemaren administratzailearen helburua aplikazioaren funtzionamendu egokia bermatzea da. Honetarako, erabiltzaile honek web aplikazioan tresna gehigarriak izango ditu atzigarri administrazio lanak burutzeko, esate baterako: erabiltzaileak eta fitxategiak kudeatzeko kontrol panela. Gainera, administratzaileak erabiltzaile aurreratuak kategoriaz igotzeko baliabideak izango ditu, era honetan, konfiantzazko erabiltzaileak sortuz. Konfiantzazko erabiltzaileen helburua fitxategiak sistemara igotzeko eta administratzaileak erabiltzaile hauen materiala errebisatzeko behar duen denbora tartea ezabatzea da.

Hortaz, webgunean testu eta ariketa desberdinak egongo dira eskuragarri erabiltzaileentzat, batzuk publikoak izango dira, edozeinentzat erabilgarri, eta beste batzuk pribatuak, hauek fitxategien jabeentzat mugatuta. Fitxategi baten jabea hau sistemara igotzeko duen erabiltzailea izango da, eta bi partekatze modu izango ditu aukeran: publikoa edo pribatua. Fitxategien partekatze modua sistemara igotzeko duen erabiltzailearen menpe egongo da, hau da, testua igotzerako garaian erabaki beharko da ea erabiltzaile guztientzat

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

atzigarri egongo den edota erabilera pribatura mugatu nahi den. Edozein kasutan jabeak bere fitxategien partekatze modua aldatzeko aukera izango du.

Erabiltzaile erregistratu batek testu fitxategi bat igotzen duen bakoitzean, galdera sortzaile sistema automatikoki exekutatu da eta testu fitxategiaren gaineko galderak sortuko ditu. Hala ere, web aplikazioan material hori guztia onartzeko erabiltzailearentzat gardena den tarteko pauso bat burutuko da. Izan ere, administratzaileak testua baliozkotu beharko du erabiltzaileak etikaz jokatzen duela egiaztatuz. Honetarako, testua irakurri beharko da eta edozein erabiltzailearentzat aproposa dela egiaztatu. Aitzitik, konfiantzazko erabiltzaileen kasuan, aipatu berri dugun tarteko pauso hau ez da beharrezkoa izango.

Ariketak lantzeko garaian posible izango da: testuen gaia edota zailtasun maila ikustea, testuaren galderak aztertzea, beste erabiltzaileek testu horri eman dioten balorazioaren batz bestekoa ikustea, iritziak irakurtzea edota landu nahi diren ariketen fenomeno linguistikoak ⁶ zein galdera kopurua zehaztea.

Balorazioa fitxategiak eta sortutako galderak ebaluatzeko irizpide bat da. Web aplikazioan erregistratuta dagoen edozein erabiltzaile izango da gai fitxategiak baloratzeko -galdera zentzudunak, interesgarriak testuaren mamia-rikiko, euskara lantzeko aproposak ote diren, ...-. Gainera, fitxategien inguruan iritziak irakurtzeko eta idazteko -azken hau erabiltzaile erregistratuentzat bakarrik- aukera ere eskainiko da.

Funtzionalitate guztiak era antolatu eta erosoan izateko, erabiltzaileek bi kudeaketa gune nagusi erabil ditzakete: ariketak lantzeko txokoa, erabiltzaile guztientzat atzigarri egongo dena, eta fitxategiak kudeatzeko txokoa, azken hau erabiltzaile erregistratuentzat mugatua. Alde batetik, fitxategiak kudeatzeko txokoak aukera emango dio sisteman atzigarri dituen testuen informazio guztia ikusteko, edota hauen gainean eragiketak burutzeko.

Testu batekin erlazionatutako informaziorik garrantzitsuena ondorengoa da: idazlanaren gaia, zailtasun maila, esleitutako balorazioa, partekatze modua, data, honi loturiko iritziak, balorazioak eta onarpena. Testuen gainean burutu daitezkeen eragiketak ondorengoak dira: dokumentua ezabatu, erlazionaturiko galdera sorta edo dokumentua bera atzitu, galdera okerrak ezabatu eta zuzendu, galdera berriak sortu, balorazioa ikusi edota eman, partekatzea aldatu, erabiltzaileek idatzi dituzten iritziak irakurri eta fitxategietan iritziak idaztea. Bestetik, ariketak lantzeko txokoak, aplikazioaren atalik garrantzitsuena izanik, QG sistemak automatikoki sortu edo/eta fitxategiaren

⁶Fenomeno linguistikoak sintagma ezberdinen deklinabide kasuak zehazteko erabiltzen dira, esate baterako: absolutiboa, ergatiboa, ablatiboa, adlatiboa, genitiboa ...

jabeak eskuz editatu duen materiala lantzeko aukera ematen du, honetarako, erabiltzailearen elkarrekintzekin interaktuatzen duen interfaze bat erabiltzen da.

Honetaz gain, ariketak lantzeko garaian galderen erantzuna zuzena den ala ez konprobatzeko aukera egongo da, eta, era honetan, ariketak automatikoki zuzentzea posible izango da. Prozesu honetan, itzulpengintzako metrika desberdinak erabili dira, eta hastapenak baino landu ez badira, esaldien urruntasuna neurtzen da hauek zuzentzat edo okertzat jotzeko. Halaber, emaitza zuzena eta galdera erantzuten lagunduko digun pista ikusteko aukera ere eskainiko da.

4.3.2 Modulu nagusien diseinua

- Webgunearen sarrera orrialdea

Erabiltzaile orok jasotzen duen interfazea, web aplikazioan esku-ragarri dauden tresnak modu antolatuan bistaritzen dituena. Erabiltzaileak aplikazioko modulu nagusienetara zuzenduko dituzten estekak integratzeaz gain, beste baliabide batzuk ere biltzen ditu; kasu erabiltzaile kontuekin loturiko hainbat baliabide, esaterako: pasahitza berreskuratu, kontu berria sortu, kautotu, kontua ezabatu, ...

- Informazio modulua

Erabiltzaile hasiberrientzat zuzendua, web aplikazioa zertan datzan azaltzen da bertan. Gainera, inplementazioari dagokion kapituluan aipatuko den *Search Engine Optimization* ataza eraginkortasunez egikaritzeko lagungarri da. Izan ere, orrialde hauetan aplikazioaren nondik norakoak laburbiltzen dira eta informazio hau oso baliagarria da web bilatzaileentzat.

- Iritziak ikusi eta idazteko modulua

Erabiltzaileek web aplikazioan atzigarri dituzten fitxategien inguruan iritziak, iruzkinak eta azalpenak irakurtzeko eta idazteko baliabidea. Modulu honen helburua erabiltzaileen arteko elkarrekintza eta elkarlana sustatzea da.

- Balorazioak ikusi eta idazteko modulua

Balorazioak fitxategiak eta hauen inguruan sortutako galderak ebaluatzea dute helburu. Balorazioak iritzien pareko badira ere, helburu desberdina dute, izan ere, balorazioak erabiltzailearen balioespena zenbaki batean laburbilduko duen bitartean, iritziek, erabiltzaileen kriti-

kak jasotzeko aukera emango dute. Gainera, fitxategiak aurkezteko garaian balorazio guztien bataz bestekoa pantailaratuko da fitxategiaren informazioarekin batera, iritziak, aldiz, ez dira zuzenean bistaratuko.

- Erabiltzaileak kudeatzeko modulua

Erabiltzaile kontuak kudeatzeko baliabideak eta oinarritzko funtzioak integratzen ditu modulu honek, funtzio eta baliabide hauetako asko administratzailearentzat mugaturik daude.

- Ariketak lantzeko modulua

Modulu honetan web aplikazioan erabilgarri diren fitxategi guztiak azalduko dira, honetaz gain, fitxategien inguruko informazioa ere azalduko da eta erabiltzailearen esanetara filtratu, landu nahi den fitxategiaren aukeraketa errazteko asmoz. Behin erabiltzaileak fitxategi bat aukeratu ondoren galderak erantzuteko modulura jauzi egingo du.

- Galderak erantzuteko modulua

Aukeratu den fitxategiarekin lanean aritzeko interfazea. Ariketak lantzeko tresna lagungarriak, hiztegiak esaterako; interfazean agertzen diren galderen inguruko informazioa eta galderak erantzuteko lan ingurunea biltzen ditu interfaze honek. Honetaz gain, sortu nahi diren ariketak egokitzea posible izango da filtroaz baliatuz, era honetan, landu nahi diren galdera motak eta galdera kopurua zehazteko aukera egongo da. Galderen emaitzak zuzentzeko itzulpen automatikoan erabiltzen diren metrika desberdinekin osaturik dagoen zuzentzailea garatu da -ikus 4.29 irudia-.

- Fitxategiak kudeatzeko modulua

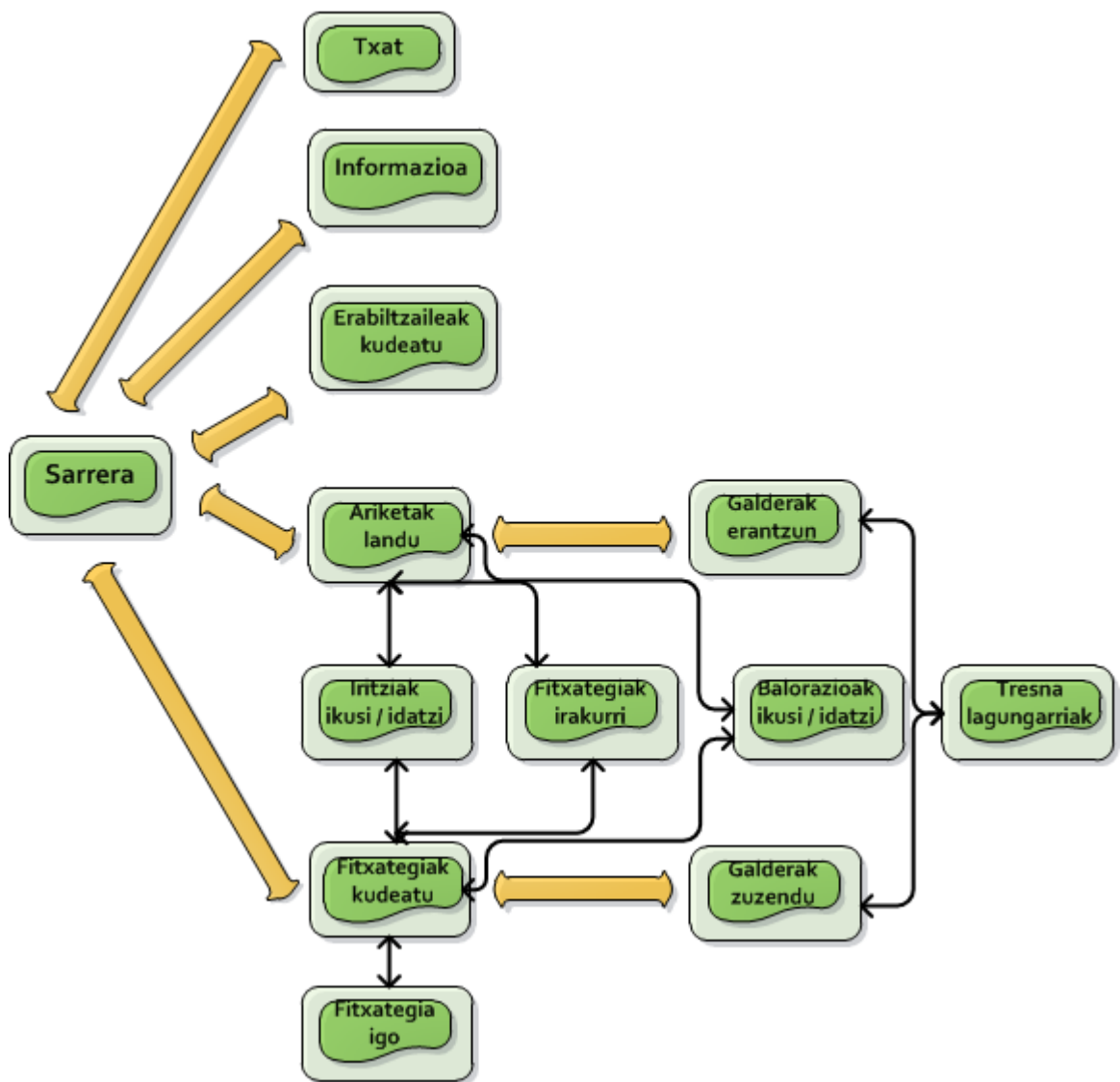
Modulu honetan web aplikazioan erabiltzaileak dituen fitxategi guztiak azalduko dira, honetaz gain, fitxategien inguruko informazioa ere azalduko da. Behin erabiltzaileak fitxategi bat aukeratu duenean galderak zuzentzeko modulura jauzi egingo du. Honetaz gain, fitxategiak aplikaziora igotzeko baliabideak ere modulu honetan integratuta aurkeztuko dira interfazean.

- Galderak zuzentzeko modulua

Modulu honek aukeratu den fitxategiaren galderak zuzentzea du helburu. Tresna lagungarriak, interfazean agertzen diren galderen inguruko informazioa eta galderak zuzentzeko lan ingurunea biltzen ditu interfaze honek. Modulu honen bitartez galdera berriak sortzea, gaizki

dauden galderak zuzentzea eta ezabatzea eta laguntza editatzea posible izango da.

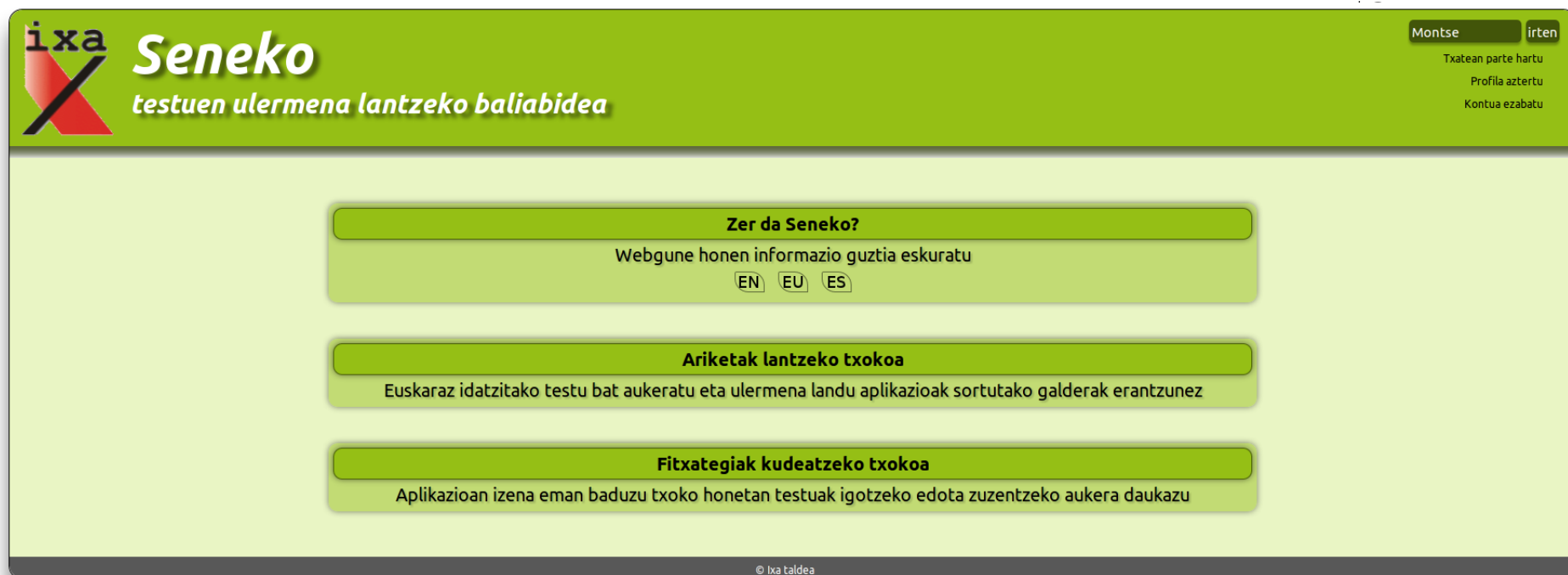
Hurrengo irudian moduluek osatzen duten egituraren eskema azaltzen da, eta, honen ostean, web aplikazioaren moduluen irudiak:



Irudia 4.3: Web aplikazioaren diseinuaren eskema.



Irudia 4.4: Sarrera orrialdearen interfazea.



Irudia 4.5: Sarrera orrialdearen interfazea erabiltzaile kautotuentzat.

idErabiltzaile	onartua	izena	korreoElektronikoa	kautotzeak (zuzen/oker)	dataAlta	Fitxategi kopurua	Konfiantzakoa
1	1			120 / 15	2012-10-05	4 (0 proz /0 onartz)	bai
7	1			12 / 1	2012-10-30	0 (0 proz /0 onartz)	bai
8	1			40 / 4	2012-12-11	10 (0 proz /0 onartz)	bai
10	1			2 / 0	2013-01-07	1 (0 proz /0 onartz)	bai
20	1			8 / 2	2013-01-08	1 (0 proz /0 onartz)	bai
21	1			4 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
22	1			4 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
23	1			2 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
24	1			5 / 0	2013-01-09	0 (0 proz /0 onartz)	bai
25	1			3 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
27	1			2 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
28	1			2 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
29	1			2 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
30	1			11 / 1	2013-01-09	3 (0 proz /0 onartz)	bai
31	1			3 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
32	1			2 / 0	2013-01-09	1 (0 proz /0 onartz)	bai
33	1			1 / 0	2013-01-09	0 (0 proz /0 onartz)	ez
37	1			0 / 0	2013-01-09	0 (0 proz /0 onartz)	ez

Irudia 4.6: Erabiltzaileak kudeatzeko interfazea.

Erabiltzaile baten datu guztiak erazi **Aukeratu erabiltzailea** **Konfiantza kendu** **Konfiantzazko bihurtu** **Erabiltzailea ezabatu**

Izena

Abizenak

Pasahitza

Posta kodea

Adina

Euskara maila

Onartze kodea

Path

Fitxategi publikoen zerrenda

idFitxategi	izena	izenaXML	zailtasunMaila	gaia	onartua	dataAlta	Irakurri	Onartu	Ezabatu
178	Azeria_eta_saia.txt	Azeria_eta_saia.txt.xml	L.H. (1.zikloa)	Ipuina	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
184	Errege_alferra.txt	Errege_alferra.txt.xml	L.H. (2.zikloa)	Ipuina	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
188	Nire_aitona_Balentin.txt	Nire_aitona_Balentin.txt.xml	L.H. (2.zikloa)	Ipuina	1	2013/01/07	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>

Fitxategi pribatuen zerrenda

idFitxategi	izena	izenaXML	zailtasunMaila	gaia	onartua	dataAlta	Irakurri	Onartu	Ezabatu
179	Euskal_Herriko_mitologia.txt	Euskal_Herriko_mitologia.txt.xml	L.H. (1.zikloa)	Mitologia	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
180	Inauteriak.txt	Inauteriak.txt.xml	L.H. (1.zikloa)	Kultura	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
181	Itsas_botika_niminoenak.txt	Itsas_botika_niminoenak.txt.xml	L.H. (1.zikloa)	Medikuntza	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
182	Lazkao_Txikiren_biografia.txt	Lazkao_Txikiren_biografia.txt.xml	L.H. (1.zikloa)	Pertsonaiak	1	2012/12/12	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
185	Frantziako_iraultza.txt	Frantziako_iraultza.txt.xml	D.B.H. (2.zikloa)	Historia	1	2012/12/26	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
186	Munduko_bigarren_gerra.txt	Munduko_bigarren_gerra.txt.xml	D.B.H. (2.zikloa)	Historia	1	2012/12/26	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>
204	Frantziako_iraultza_2.txt	Frantziako_iraultza_2.txt.xml	D.B.H. (2.zikloa)	Historia	1	2013/01/10	<input type="button" value="Irakurri"/>	<input type="button" value="Onartu"/>	<input type="button" value="Ezabatu"/>

© Ixa taldea Webgunearen hasierako orrira itzuli

Irudia 4.7: Erabiltzaileak kudeatzeko interfazea.

Ariketak lantzeko txokoa

Fitxategi publikoen zerrenda - Ezkutatu
Emailzak filtratu
Gaia - edozein
Zailtasun maila - edozein
Balorazioa - edozein
Egilea - edozein

Izena	Gaia	Zailtasun maila	Balorazioa	Data	Egilea		
Azeria eta saia.txt	Ipuina	L.H. (1.zikloa)	8 Iritziak ikusi	2012/12/12		Irakurri	Ariketak landu
Errege alferra.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	2012/12/12		Irakurri	Ariketak landu
Nire aitona Balentin.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	2013/01/07		Irakurri	Ariketak landu
gailurrak.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/09		Irakurri	Ariketak landu
parrandatxo.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/09		Irakurri	Ariketak landu
Manexen ipuina.txt	Ipuina	L.H. (3.zikloa)	baloratu gabe Iritziak ikusi	2013/01/10		Irakurri	Ariketak landu
fitxategia Olatz ipuina lh2.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	2013/01/10		Irakurri	Ariketak landu
aitona baserritarra.txt	Antzerkia	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	2013/01/10		Irakurri	Ariketak landu
NIREaitona.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	2013/01/10		Irakurri	Ariketak landu
sorginapirulinak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
katuegalariak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
dantzariak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
lapurbihurriak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
lotiederrak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
baserritarrak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
ausartak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu
txakurabeslariak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	2013/01/13		Irakurri	Ariketak landu

Nire fitxategien zerrenda - Ezkutatu

Izena	Gaia	Zailtasun maila	Balorazioa	Onartua	Data		
Azeria eta saia.txt	Ipuina	L.H. (1.zikloa)	8 Iritziak ikusi	bai	2012/12/12	Irakurri	Ariketak landu
Euskal Herriko mitologia.txt	Mitologia	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	bai	2012/12/12	Irakurri	Ariketak landu
Inauteriak.txt	Kultura	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	bai	2012/12/12	Irakurri	Ariketak landu

Irudia 4.8: Ariketak lantzeko interfazea.



Montse irten
Txatean parte hartu
Profila azertu
Kontua ezabatu

Fitxategiak kudeatzeko txokoa

Nire fitxategien zerrenda - Ezkutatu

Izena	Gala	Zailtasun maila	Balorazioa	Partekatzea	Data	Onartua			
Azeria eta saia.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	publikoa aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Euskal Herriko mitologia.txt	Mitologia	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Inauteriak.txt	Kultura	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Itsas botika niminoenak.txt	Medikuntza	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Lazkao Txikiren biografia.txt	Pertsonaiak	L.H. (1.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Errege alferra.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	publikoa aldatu	2012/12/12	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Frantziako iraultza.txt	Historia	D.B.H. (2.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/26	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Munduko bigarren gerra.txt	Historia	D.B.H. (2.zikloa)	baloratu gabe Iritziak ikusi	pribatua aldatu	2012/12/26	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
Nire aitona Balentin.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe Iritziak ikusi	publikoa aldatu	2013/01/07	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu

Fitxategi bat aplikazlora igo

Izena

Gala Antzerkia ▾

Zailtasun maila L.H. (1.zikloa) ▾

Partekatze modua pribatua ▾

Fitxategia aukeratu (txt formatua, gehienez 25 kilo byte) Examinar...

Era **etikoan** jokatuko dut Igo

© Ixa taldea Webgunearen hasierako orrira itzuli

Irudia 4.9: Fitxategiak kudeatzeko interfazea.

Galderak kudeatzeko txokoa

Fitxategia : Azerie eta saia.txt

Hautaturiko fitxategiari buruzko ariketen informazioa Ezkutatu

Galdera kopurua guztira:	2
Galdera motak :	
Nor / Zer :	1
Norengandik / Nondik :	0
Norenganaino / Noraino :	0
Norantz :	0
Norengana / Nora :	0
Norekin / Zerekin	0
Nori / Zeri :	0
Norentzat / Zerentzat :	0
Nork / Zerk :	0
Noren / Nongo :	0
Noren / Zeren :	0
Norengan / Non :	1

Tresna lagungarriak

- Elhuyar itzultzailea
- Euskal Hiztegia
- Hiztegi Batua
- Hiztegi gehiago oraindik!
- >Irakurri testua leiho berrian<--

Galdera	Zer erantzun zion azeriak?	Ezabatu galdera
Erantzuna	Hemen nago goitik behera datozen horiek zer ote diren begira	
Laguntza	Hemen nago goitik behera datozen horiek zer ote diren begira erantzun zion azeriak	
Kasua	Absolutiboa	
Galdera	Non zegoen azeria?	Ezabatu galdera
Erantzuna	Hemen	
Laguntza	Hemen nago erantzun zion azeriak	
Kasua	Inesiboa	

Galdera berria sortu	Galdera berri bat sortu nahi baduzu, aukeratu kasua eta sakatu botoia	Nor / Zer	Sortu galdera
Galderak gorde	Galderak zuzendu ondoren sisteman zuzenketak gorde nahi badituzu, sakatu ondorengo botoia (aurreko edukia galduko da)	gorde	
Berreskuratu	Akats bat egin baduzu, ez urduritu, ondorengo botoia sakatuz aurreko egoera berreskuratuko duzu	berreskuratu	

© Ixa taldea Webgunearen hasierako orrira itzuli

Irudia 4.10: Galderak zuzentzeko -editatzeko- interfazea.

Ariketak lantzeko txokoa

Fitxategia : Nire aitona Balentin.txt

Hautaturiko fitxategiari buruzko ariketen informazioa Erakutsi

1. Galdera Nora jaitsi zen aurreko asteazkenean?

Erantzuna ✗

Erantzun egokia

Laguntza

2. Galdera Nortzuk hasi ziren barrezka?

Erantzuna ✓

Erantzun egokia ezkutatu

Laguntza

3. Galdera Noiz jaitsi zen herrira?

Erantzuna

Erantzun egokia Zuzendu

Laguntza ikusi

4. Galdera Zer da nire aitona?

Erantzuna

Erantzun egokia

Laguntza

Tresna lagungarriak

- Elhuyar hiztegia
- Orotariko Euskal Hiztegia
- Hiztegi Batua
- Bestelako hiztegiak
- Laguntza desaktibatu
- >Irakurri testua leiho berrian<--

© Ixa taldea Webgunearen hasierako orrira itzuli

Irudia 4.11: Galderak erantzuteko -lantzeko- interfazea.

4.4 Galdera sortzaile eta web aplikazioaren arteko komunikazioa

Proiektu honetan garatu den web aplikazio eta galdera sortzaile sistemaren arteko informazio trukea bermatzeko, galdera sortzaile sistemaren lehen bertsioan ikusitako sarrera eta irteera aldatzea beharrezkoa izan da. Izan ere, komunikazio eredu berrian, web aplikazioak analizatu behar du erabiltzaileak sistemara igo duen sarrera corpusa, eta, ondoren, informazio linguistikoa bil-tzen duen fitxategi hau galdera sortzaile sistemari bidali behar dio. Era be-rean, automatikoki sortutako galderen fitxategia web aplikazioan biltegiratu behar da.

Komunikazio eredu berri hau bi prozesu sekuentzialetan banatu da: alde batetik, sarrera corpusa analitzeko prozesua, eta bestetik, galderak sortze-ko prozesua. Ondorengo lerroetan aurkeztutako bi prozesu hauek sakonago azaltzen dira.

4.4.1 Sarrera corpusa analitzatzeko prozesua

Testu fitxategiak linguistikoki analizatzeaz eta emaitzak fitxategi sisteman biltegiratzeaz arduratzen da prozesu hau; bertan erabiltzaileak, fitxategiak eta analizatzaile linguistikoak parte hartzen dute. Ondoren, prozesuak ja-rraitzen dituen pausoak laburbiltzen dira:

1. Erabiltzaileak testu fitxategia gordetzen du webgunean.

Dokumentua fitxategi sisteman gordetzen den bezain azkar, hu-rrengo atazak martxan jarriko dituen azpi prozesu konkurrente bat abiarazten da. Une honetatik aurrera eta ataza hauek guztiak amaitu arte, fitxategia *prozesatzen* egoeran egongo da.

2. Fitxategia analizatzaile bati bidaltzen zaio.

Galdera sortzailearen hasierako eskeman analizatzaile gisa *txt2kaf_eu* perl aplikazioa erabiltzen bazen ere, eskema berrian web zerbitzu⁷ gisa eskainitako analizatzaile bat erabiltzen da. Besteak beste, bi arrazoi na-gusi direla eta egin da aldaketa hau: lehenengoa, aipaturiko *txt2kaf_eu* aplikazioa web zerbitzaritik zuzenean atzitu ezin delako -sixx makine-tan bakarrik erabilgarri da-, eta bigarrena, komunikazio eskema logikoki banatua izateko.

Fitxategia perl LWP moduluaren laguntzaz[9] HTTP-POST eskaera batean erantsi eta bidaltzen da web analizatzaileara.

⁷http://sixx03.si.ehu.es:8888/pipet_txt2kaf_eu

3. Fitxategiaren analisia burutzen da.

Eskaera bidali ostean prozesua zain geratzen da analisia bukatzen den arte. Arazoak ekiditeko LWP moduluan lehenetsitako *time out* denborak aldatu behar izan dira; izan ere, analisiak denbora tarte nahiko luzea har dezake, eta aldatu ostean, onarturiko denbora mugak gainditzen dira.

4. Emaitzak biltegitratzen dira.

4.4.2 Galderak sortzeko prozesua

Galderak sortzeko prozesua eta aurreko atalean aztertutakoa -corpusa linguistikoki analizatzea alegia- modu sekuentzialean exekutatzen dira; hau da, galderen sorkuntza jatorrizko fitxategiaren analisia amaitu ostean hasten da. Prozesu hau analisi linguistikoen fitxategia hartu eta galdera sortzailea Perl aplikazioa exekutatzeaz arduratzen da. Honetaz gain, emaitzak fitxategi sisteman biltegitratzen ditu, eta azkenik, webguneari prozesuaren amaieraren berri ematen dio. Prozesua laburbiltzen duen eskema hau da:

1. Analisi linguistikoen fitxategia eskuratu eta galdera sortzailea exekutatzen da.
2. Emaitzak biltegitratzen dira.

Hasieran garatutako galdera sortzailean emaitzak irteera estandarrean inprimatzen baziren ere; nabarmena da moduluen arteko informazio trukea egiteko irtenbide aproposa ez dela. Hargatik, ariketen sorkuntza helburu duten beste tresna batzuek eman dituzten urratsak jarraituz[12], galderen formatua zehaztuko duten XML eskemak diseinatu dira. Eskema hauek galderak zuzen eta modu errazean prozesatzeko aukera ematen dute. Bi eskema desberdin sortu dira: XML(1) eta XML(2).

- XML(1):

Galderen zehaztapena finkatzen duen eskema. Eskema hau galdera sortzailearen irteera da, eta horregatik, informazio linguistikoko handia biltzen du. Galderen ebaluazioari begira informazio garrantzitsua da, aldiz, webgunearentzat balio handirik ez duena. Horregatik, bigarren XML eskema bat zehaztu da.

- XML(2):

Webguneak galderekin lan egiteko finkatu den bigarren XML eskema da, lehenengoa baino askoz ere arinagoa, beharrezkoa den informazioa bakarrik biltzen duena; eta ondorioz, lan egiteko erraztasunak ematen dituena. Edozein kasutan, bi eskemak bateragarriak dira. Izan ere, XML(2) eskema XML(1) eskemak eskaintzen duen informazio guztiaren azpimultzo bat da.

3. Prozesuaren amaieraren berri ematen da.

Behin emaitzak biltzen dituen XML fitxategia biltegitatuta, azpi prozesua amaitu eta Perl LWP moduluaren laguntzaz, HTTP-POST eskaera bat bidaltzen dio webguneari amaiera honen berri izan dezan. Mezua jasotzean webguneak fitxategiaren egoera eguneratuko du.

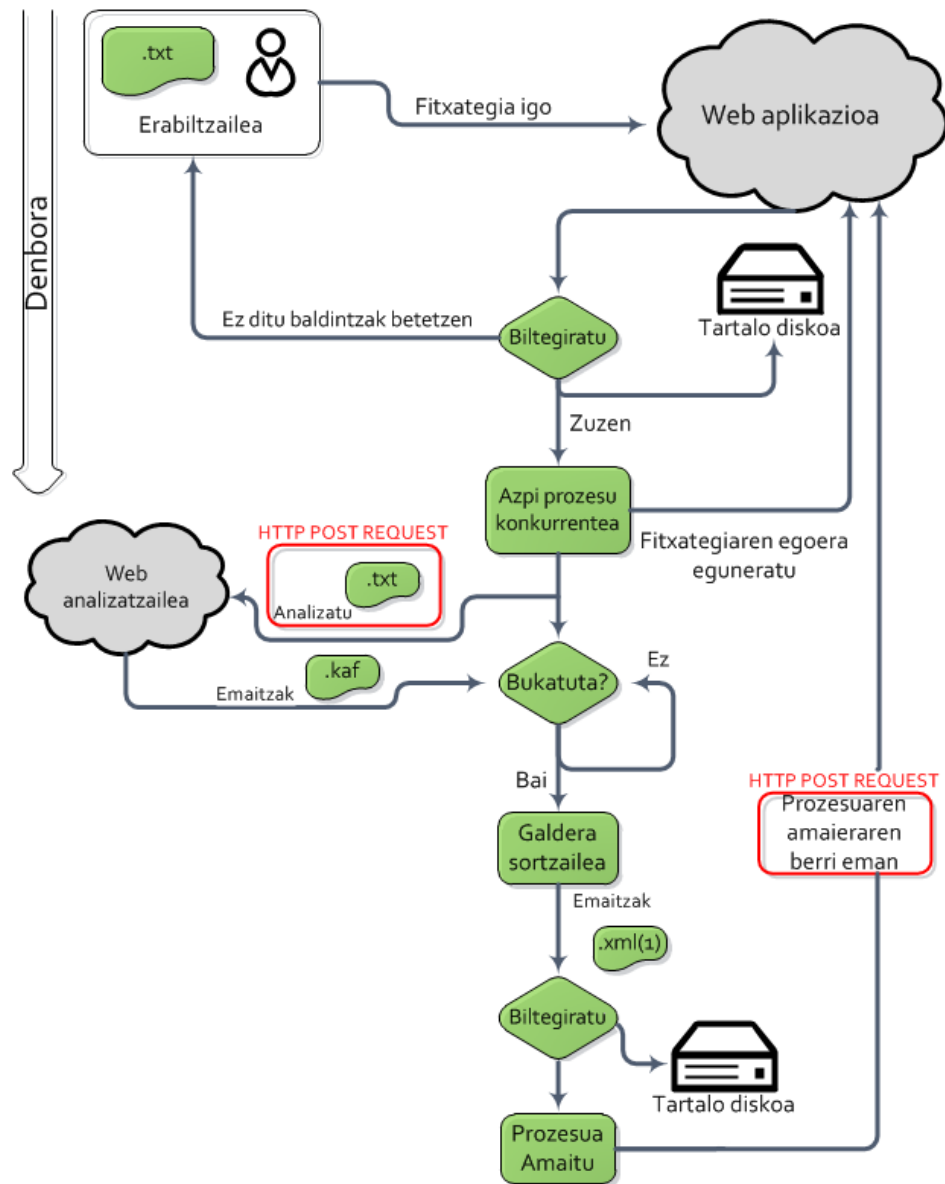
XML fitxategien zehaztapenak

XML(1) eskemaren zehaztapena

```
<?xml version="1.0" encoding="utf-8"?>
<galderak jatorri_fitxategia="">
  <galdera kasua="" heuristikoa="" jatorrizko_esaldi_id="">
    <sortutako_galdera > </sortutako_galdera >
    <jatorrizko_esaldia > </jatorrizko_esaldia >
    <galderaren_emaitza garrantzia=""> </galderaren_emaitza >
    <aditz_nagusia chunka=""
      hautagaiarekiko_distantzia_chunketan=""> </aditz_nagusia >
  </galdera >
</galderak >
```

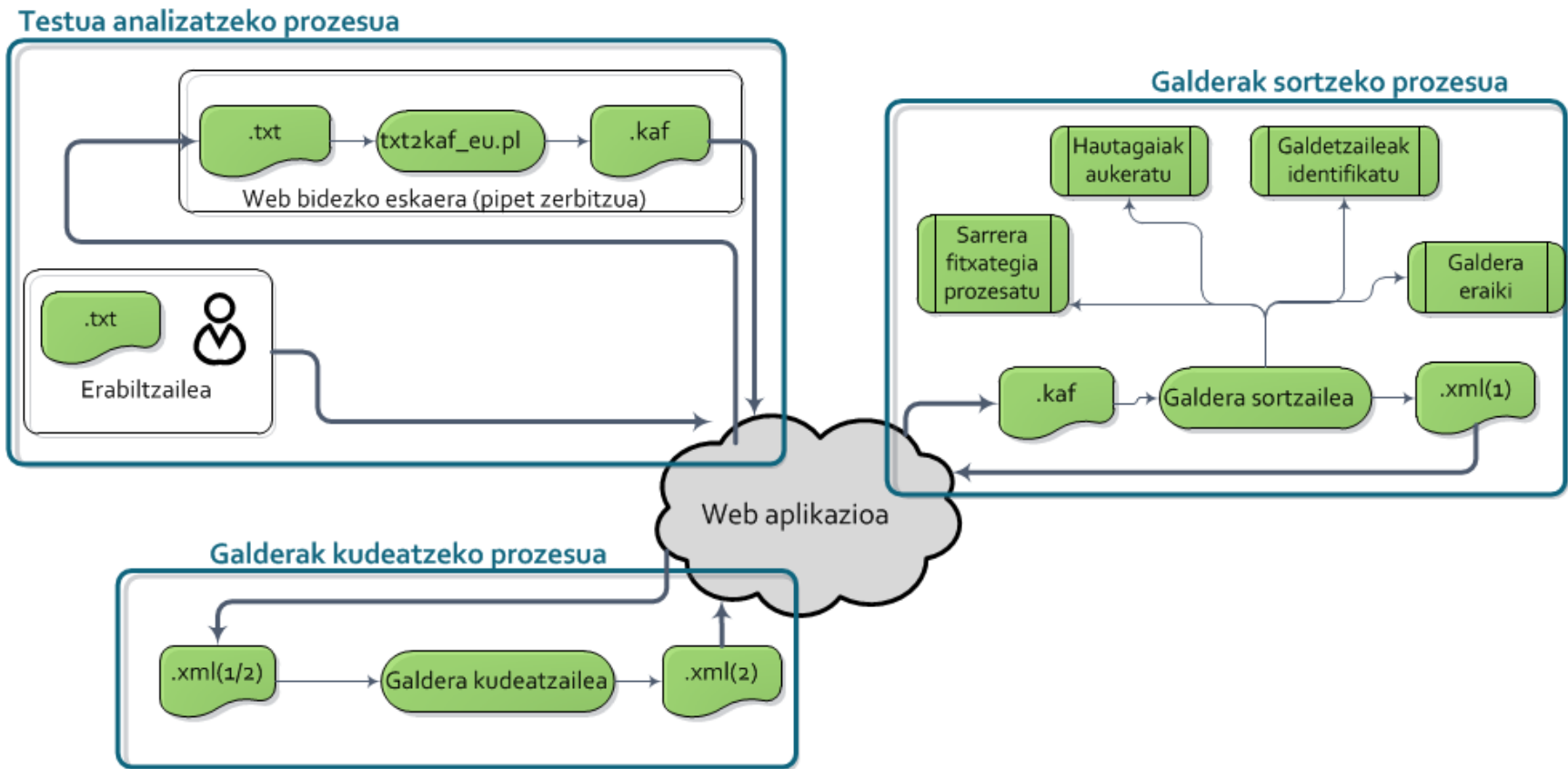
XML(2) eskemaren zehaztapena

```
<?xml version="1.0" encoding="utf-8"?>
<galderak >
  <galdera kasua="">
    <sortutako_galdera > </sortutako_galdera >
    <jatorrizko_esaldia > </jatorrizko_esaldia >
    <galderaren_emaitza > </galderaren_emaitza >
  </galdera >
</galderak >
```



Irudia 4.12: Web aplikazioak corpusak analizatzeko eta galderak sortzeko jarraitzen duen fluxu diagrama.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa



Irudia 4.13: Web aplikazio, erabiltzaile, analizatzaile sintaktiko eta galdera sortzaile sistemaren arteko komunikazio eskema.

4.5 Informazioa egituratzea eta datu basearen diseinua

Aplikazioa diseinatzeko orduan analisisian identifikatu diren atal guztiak antolatu behar dira. Diseinua objektuei zuzendutako metodologian oinarritzen da, honek abantaila eta erraztasun asko eskaintzen baititu. Ondoren zerrendatzen dira sistemaren klase nagusiak:

- Erabiltzaileak

Aplikazioko elementurik garrantzitsuenetakoak dira erabiltzaileak, sistemako beste klase guztiakin elkarrekintzak izango baitituzte. Erabiltzaileei ondoko informazioa eskatzen zaie:

- Izena
- Pasahitza
- Abizenak
- Adina
- Posta kodea
- Korreo elektronikoa
- Erregistratu den data
- Euskara maila
- Kautotze arrakastatsuen eta akastunen kopurua:

Segurtasuna zaintzea helburu duen atributua da hau, zehazki pasahitzen aurkako erasoak antzemateko.

- Konfiantzazkoa:

Administratzaileak erabiltzaile bereziak zehazteko ahalmena dauka atributu honen bitartez. Erabiltzaile bereziek abantaila nabari bat dute, administratzaileak ez baititu beren fitxategiak onartu behar; era honetan, helburua fitxategiak igotzeko beharrezko denbora tartea laburtzea da.

- Path:

Atributu honek fitxategiak non biltegitratuko diren adierazten du.

- Onartua:

Erabiltzaileak kontua gaitu duen ala ez adieraziko du atributu honek; kontua gaitzearen helburua korreo elektronikoa egiaztatzea da.

- Onartze kodea:

Erabiltzaile kontu bakoitzarekin lotzen den ausazko kodea. Erabiltzaileak kudeatzeko zein segurtasun kontuetarako erabiltzen da atributu hau.

- Fitxategiak

Klase honek erabiltzaileek sistemara igotako fitxategiak bilduko ditu. Kontuan izan beharrekoa da fitxategiak fisikoki fitxategi sisteman gordeko direla, eta ez datu basean. Hala ere, edozein unetan datu baseko fitxategi guztiak fitxategi sisteman gordetzen diren dokumentuekin sinkronizatuta egongo dira.

- Jatorrizko testuaren izena
- Testuaren inguruko galderak biltzen dituen fitxategiaren izena
- Zailtasun maila
- Gaia
- Onartua

Fitxategia erabiltzaile guztientzako aproposa dela ziurtatzeko erabiliko da atributu hau. Ziurtapen hau egiteko administratzaileak dokumentua irakurri beharko du. Honenbestez, fitxategia ezin izango da atzitu -nahiz eta publikoa izan- onartzen den unera arte. Gogoratu, aurrean esandako modura, konfiantzazko erabiltzaileen fitxategiak automatikoki onartuak izango direla.

- Partekatze mota

Partekatzea edozein kasutan publikoa ala pribatua izango da. Fitxategi publikoak edozein erabiltzailek atzitu ahal izango ditu; eta aldiz, pribatuak, fitxategiaren egileak bakarrik.

- Fitxategia sistemara igo den data

- Galderak

Galderek QG sistemak automatikoki sortutako zein erabiltzaileek sortutako, editatutako edo zuzendutako materiala bilduko dute. Galderak ez dira datu basean gordeko, aitzitik, zerbitzarietan gordeko dira XML markatze lengoia erabiliz. Galderek izango duten formatua *Galdera sortzaile eta web aplikazioaren arteko komunikazioa* atalean aipatzen da kapitulu honetan bertan (4.4).

- Iritziak

Klase honek fitxategien inguruan idazten diren iritziak bilduko ditu.

- Testu katea
- UTC⁸
- Data
- Balorazioak
 - Klase honek fitxategien kalifikazioak bilduko ditu.
 - Esleitutako balorazioa
- Kalitate kontrola
 - Klase honek ebaluazioan zehar lagungarriak izango diren atributuak biltzen ditu.
 - Hasierako galdera kopurua
 - Ezabatu diren galdera kopurua
- Txateko mezuak
 - Klase honek txat elkarreragilean idatzitako mezu guztiak bilduko ditu.
 - Testu katea
 - UTC
 - Data

Klaseak azaldu ostean, hurrengo pausoa datu base sistema diseinatzea da. Datu base sistema erabiltzea beharrezkoa da ezen web aplikazioak informazio euskarri sendoa izateko. Gainera, eskaintzen dituen ezaugarriak oso baliagarriak izango dira:

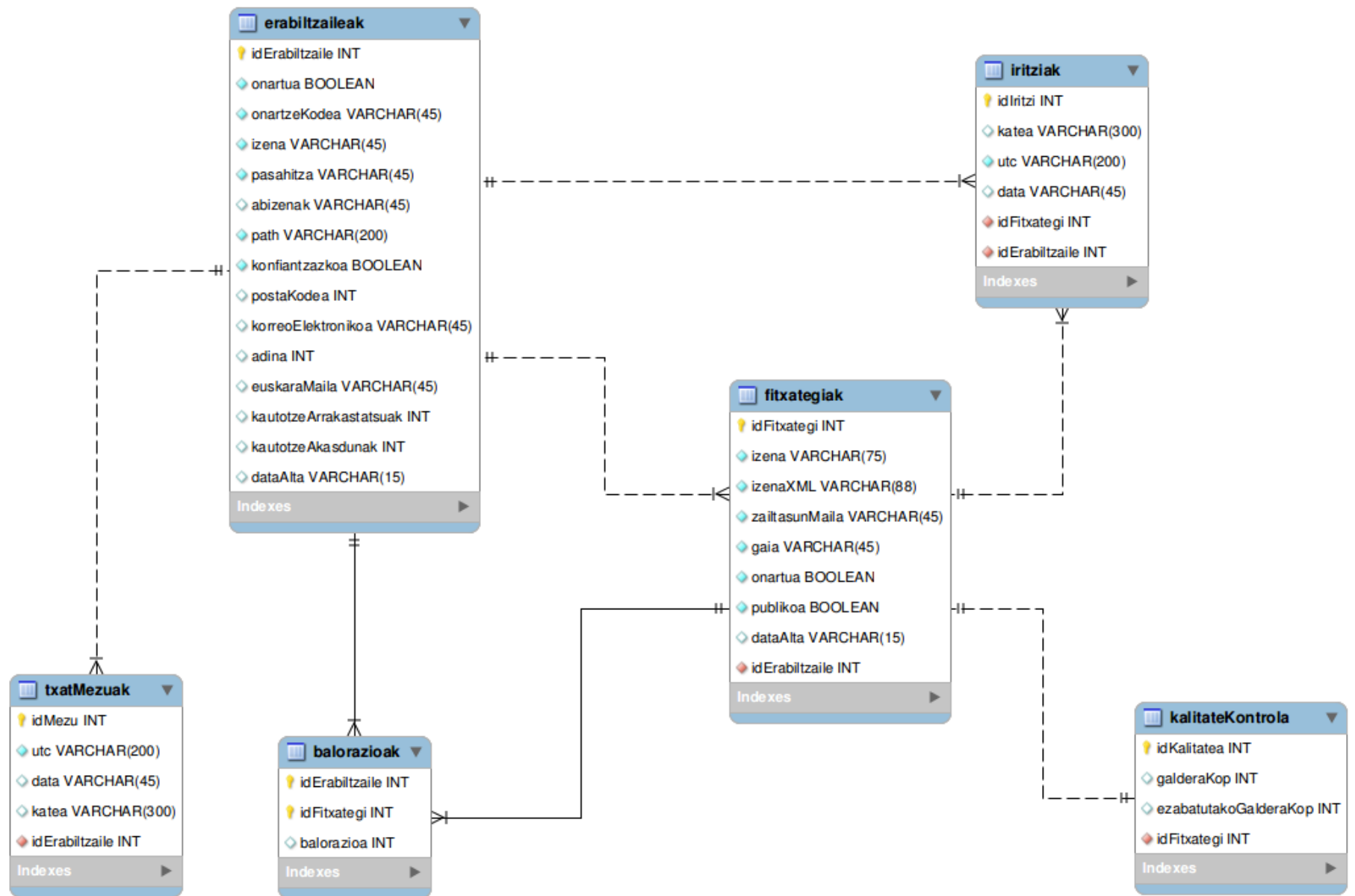
1. Informazioa atzitzeko aldiberekotasuna.
2. Informazioa modu iraunkorrean biltegitzea.
3. Informazioa atzitzeko segurtasun neurriak ezartzea.

Erabiliko den datu basea erlazionala da, hots, elementu bakoitza taula batekin identifikatzen da. Honela, elementu bakoitzak dituen eremuen arabera taulako atributuak definitzen dira eta ondoren, elementuen arteko elkarrekintzak erlazioen bidez adierazten dira.

Ondorengo irudian ikus daiteke datu base sisteman implementatuko den entitate erlazio diagrama:

⁸http://en.wikipedia.org/wiki/Coordinated_Universal_Time

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

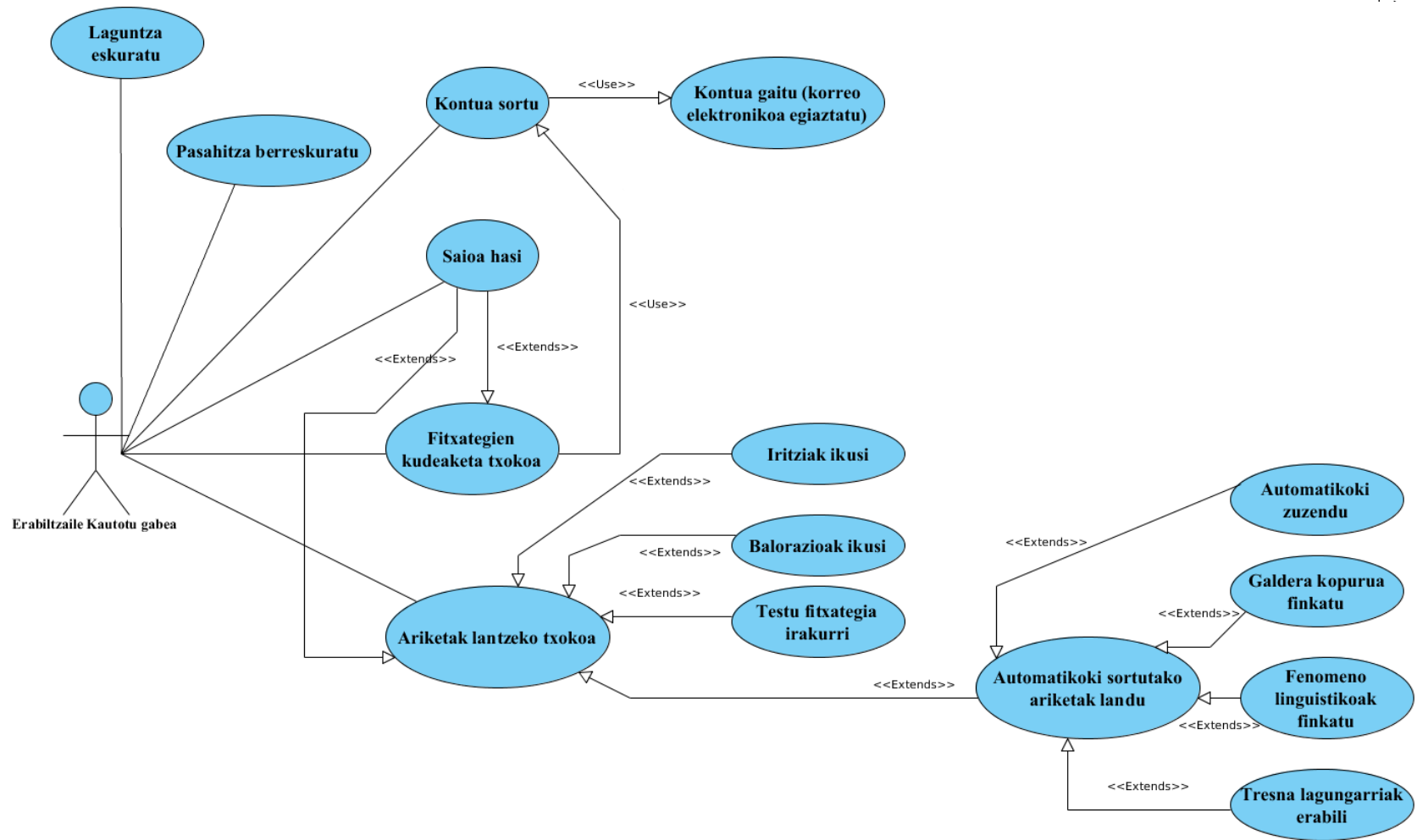


Irudia 4.14: datu baseko entitate erlazio diagrama.

4.6 Erabilpen kasuen diseinua

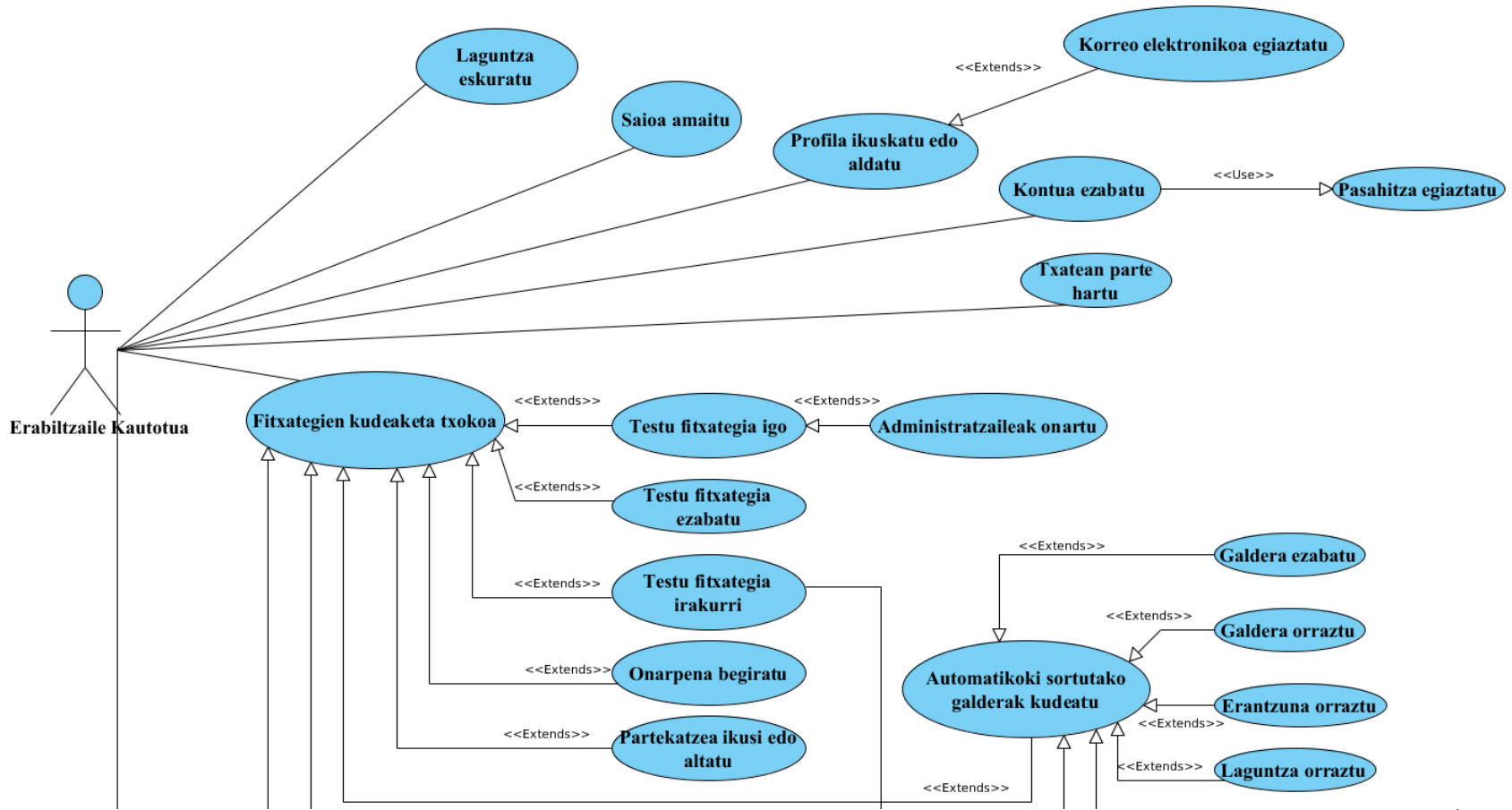
Atal honetan erabiltzaile mota guztiek izango dituzten funtzionalitateak zehaztuko dira.

4.6.1 Erabiltzaile kautotu gabeen erabilpen kasuen diagrama

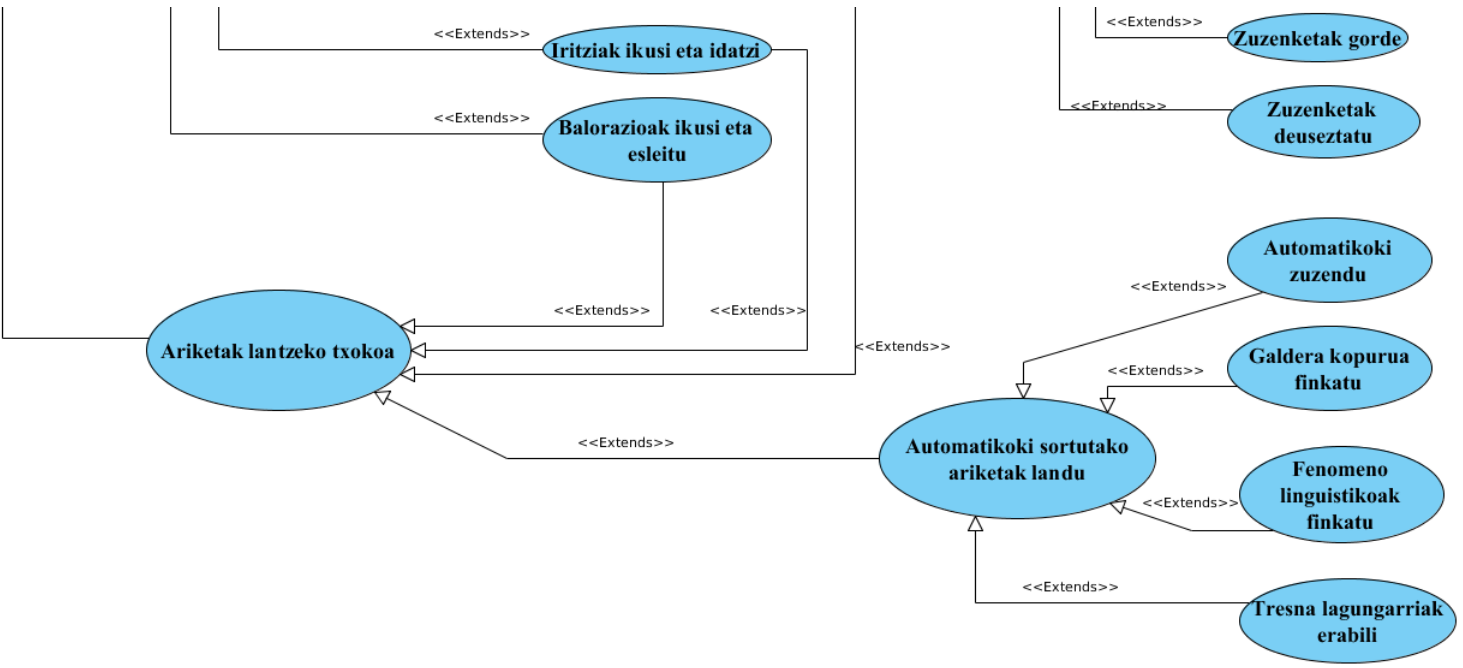


Irudia 4.15: Erabiltzaile kautotu gabeko erabilpen kasuen diagrama.

4.6.2 Erabiltzaile kautotuen erabilpen kasuen diagrama

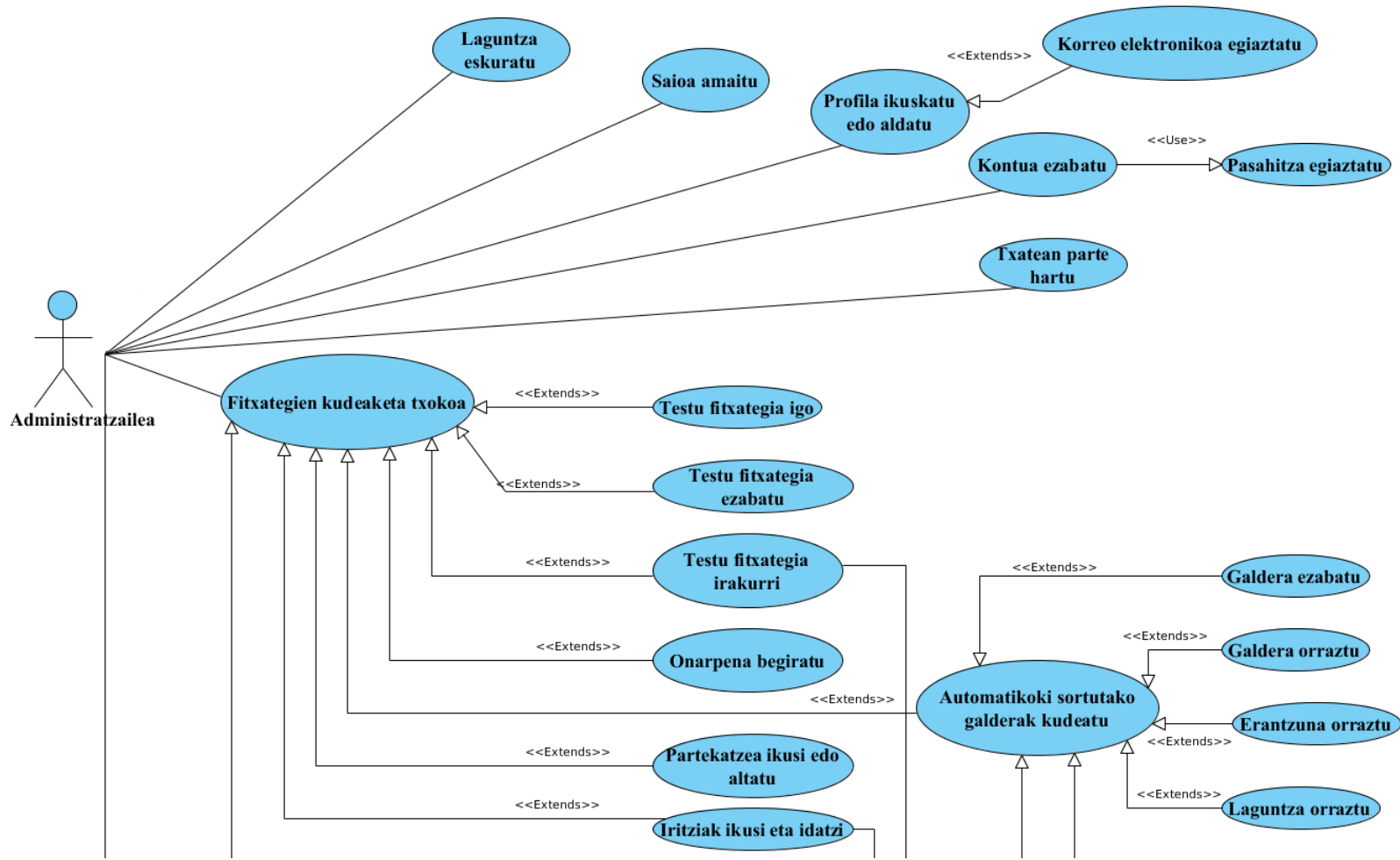


Irudia 4.16: Erabiltzaile kautotuen erabilpen kasuen diagrama.

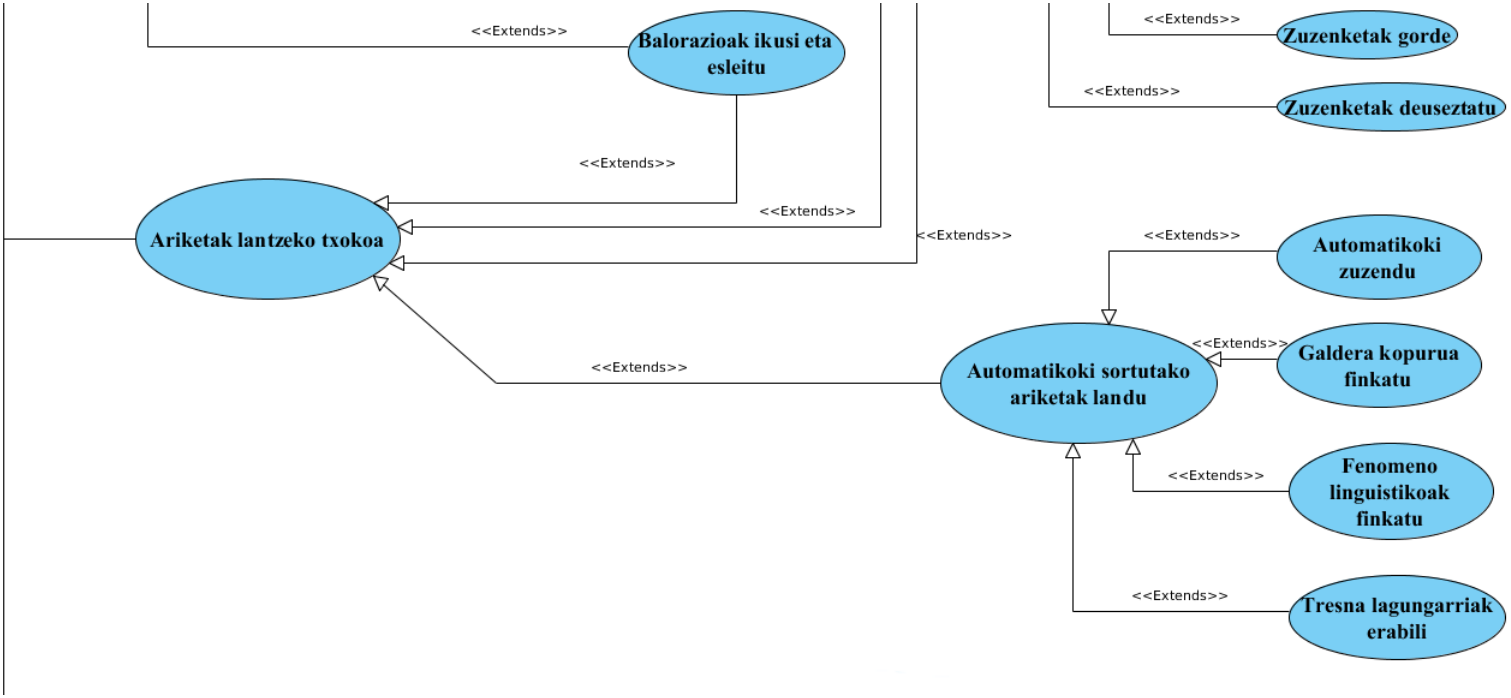


Irudia 4.17: Erabiltzaile kautotuen erabilpen kasuen diagrama.

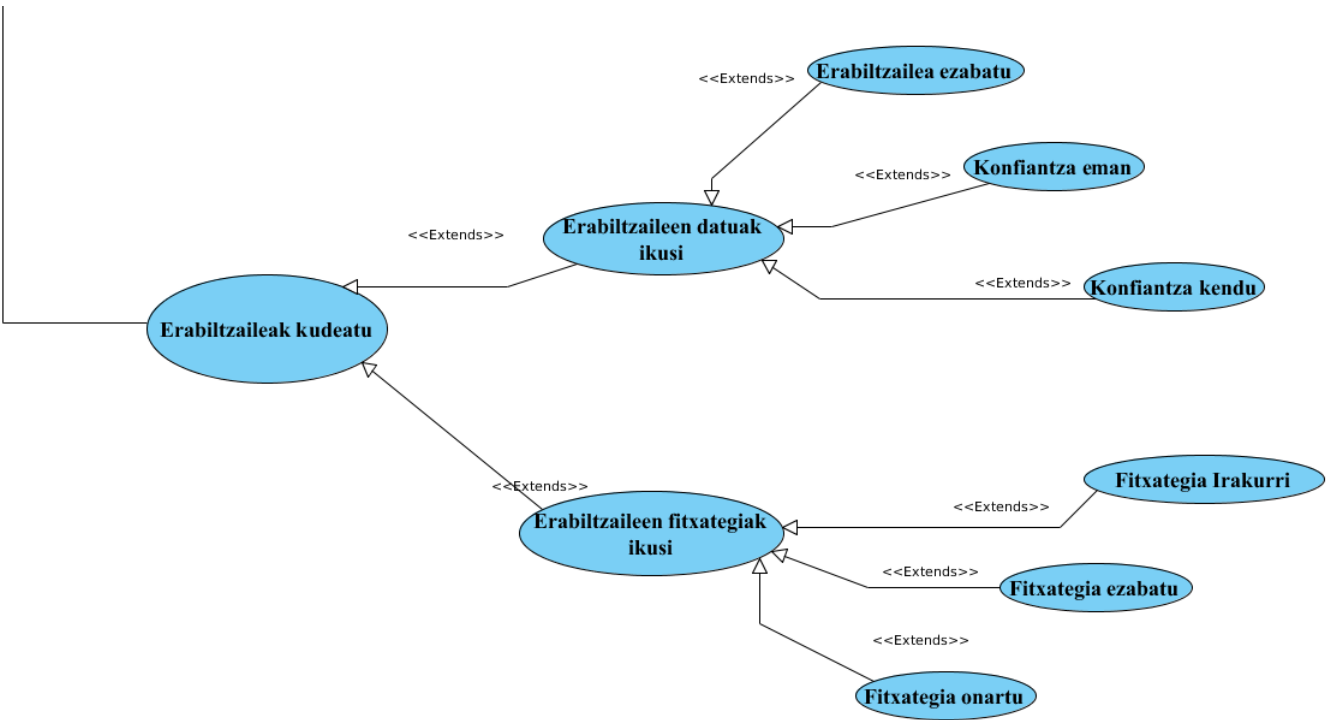
4.6.3 Administratzailearen erabilpen kasuen diagrama



Irudia 4.18: Administratzailearen erabilpen kasuen diagrama.



Irudia 4.19: Administrazioaren erabilpen kasuen diagrama.



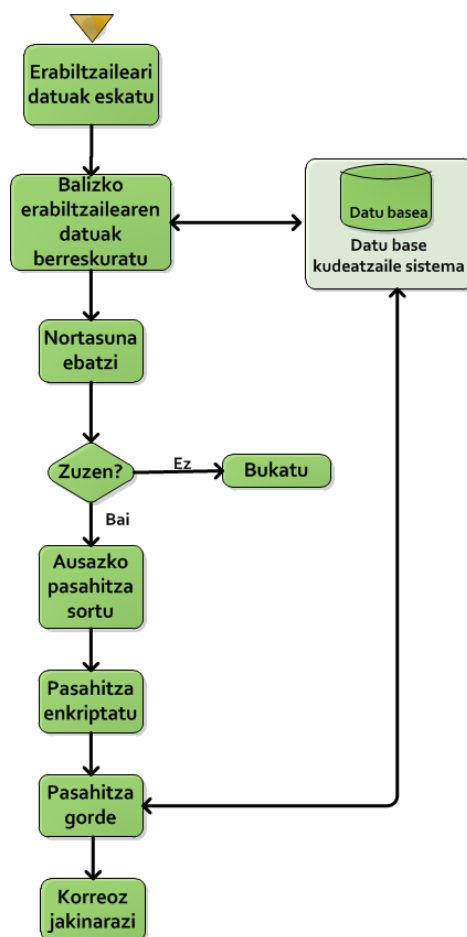
Irudia 4.20: Administratzailearen erabilpen kasuen diagrama.

4.6.4 Erabilpen kasuen fluxu diagramak

Atal honetan aipaturiko funtzionalitate garrantzitsuenak deskribatuko dituzten eskemak azalduko dira. Fluxu diagrama hauek berebiziko garrantzia izango dute aplikazioa inplementatzeari begira; funtzionalitate bakoitzaren nondik norakoak zehaztuko baitituzte.

Pasahitza berreskuratu

Erabiltzaile mota guztientzat erabilgarri, erabilpen kasu honek galdutako pasahitza berreskuratzeko aukera ematen du.

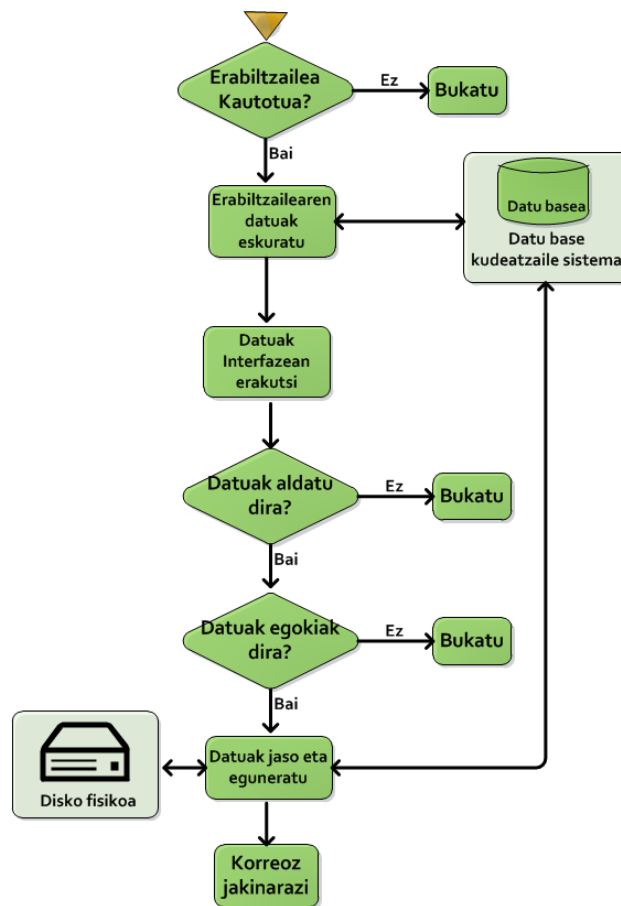


Irudia 4.21: Pasahitza berreskuratu erabilpen kasuaren fluxu diagrama.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Profila ikuskatu edo aldatu

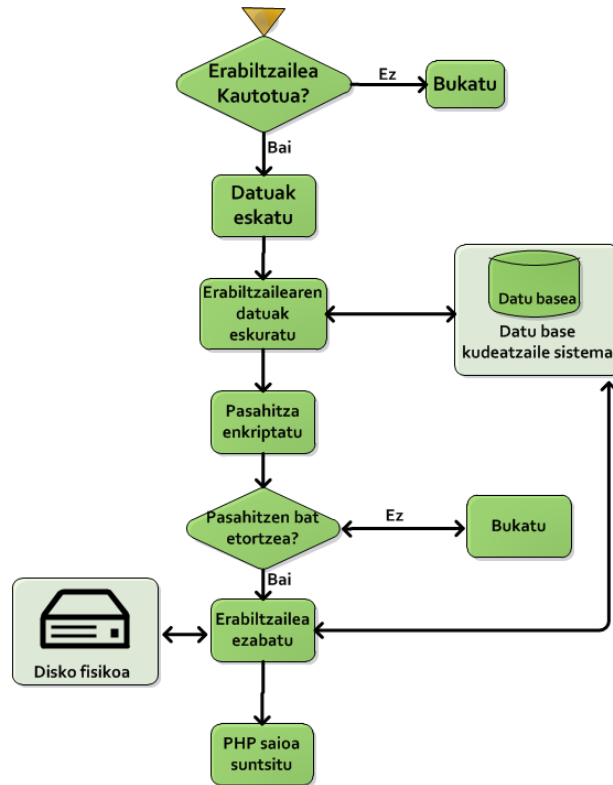
Erabiltzailea webgunean erregistratzen den garaian sortu daitezken akatsak konpontzekotan, erabilpen kasu honek edozein unetan profileko datuak aldatzeko aukera ematen du.



Irudia 4.22: Profila ikuskatu edo aldatu erabilpen kasuaren fluxu diagrama.

Kontua ezabatu

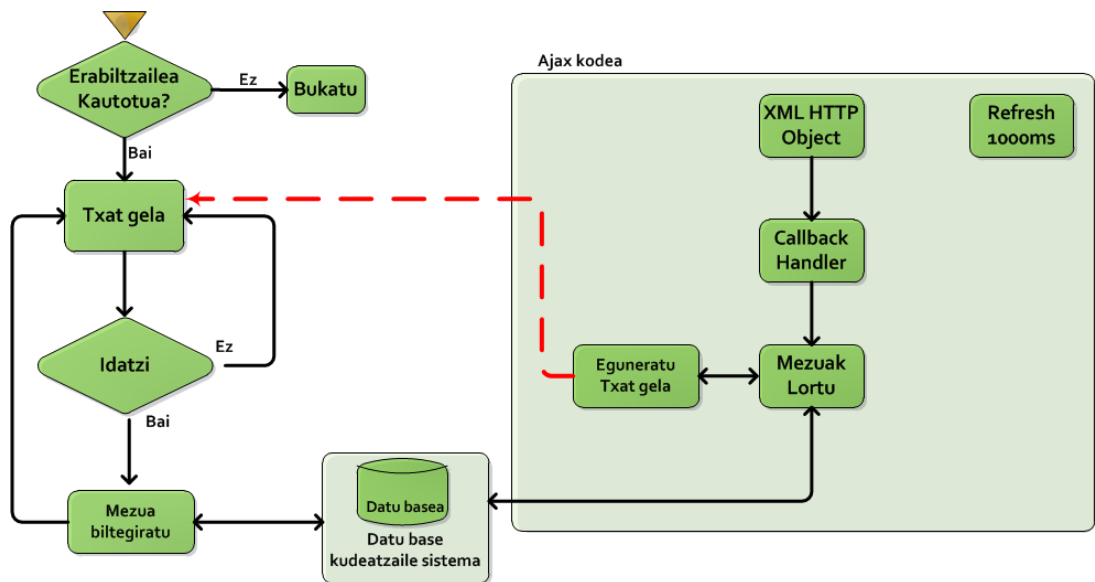
Erabilpen kasu honek erabiltzailea guztiz ezabatuko du webgunetik.



Irudia 4.23: Kontua ezabatu erabilpen kasuaren fluxu diagrama.

Txatean parte hartu

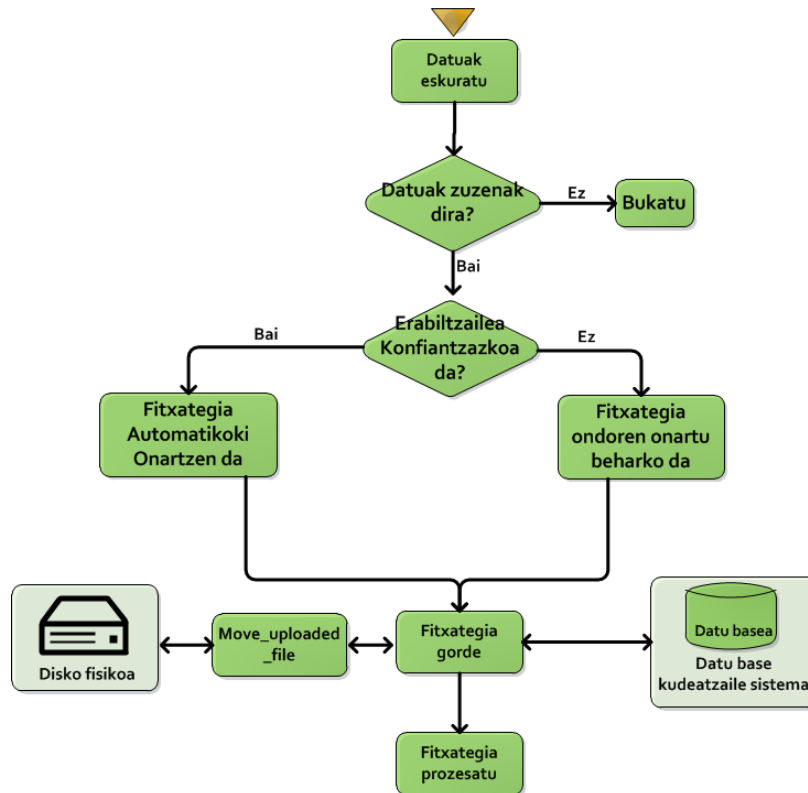
Denbora errealean erabiltzaileen arteko komunikazioa egikaritzeko aukera ematen du erabilpen kasu honek.



Irudia 4.24: Txatean parte hartu erabilpen kasuaren fluxu diagrama.

Testu fitxategia igo

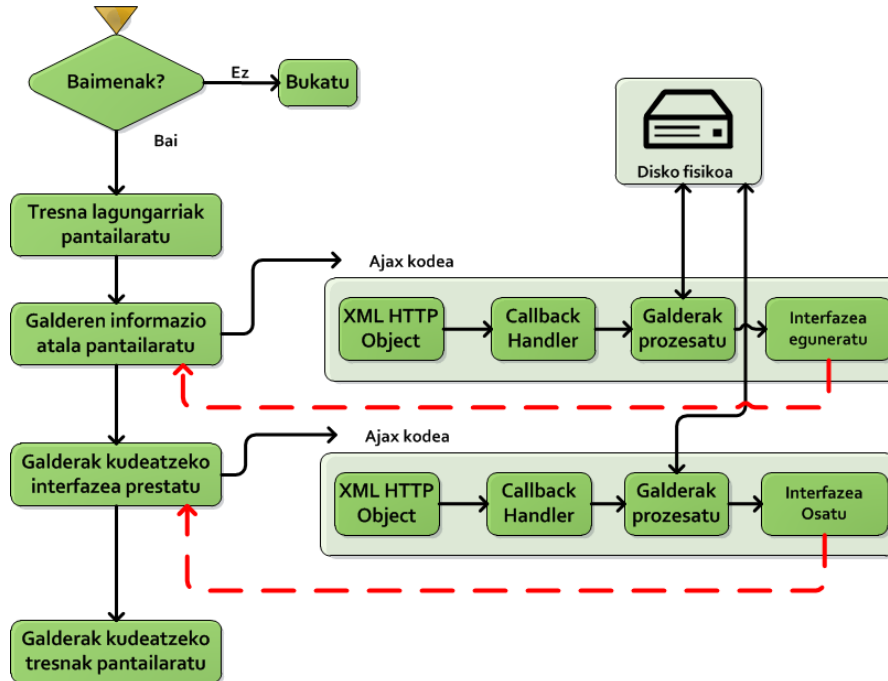
Erabilpen kasu honek erabiltzaile kautotuei webgunean fitxategiak gordetze-ko aukera ematen die.



Irudia 4.25: Testu fitxategia igo erabilpen kasuaren fluxu diagrama.

Galderak kudeatu

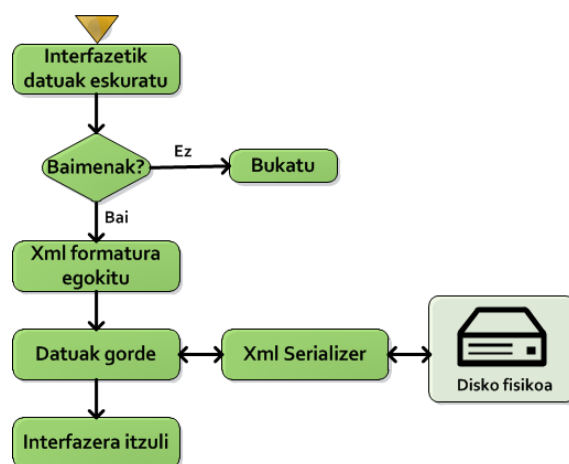
Fitxategi baten inguruan sortutako galdera, erantzuna eta laguntza aldatzeko -edo ezabatzeko- funtzionalitateak biltzen dituen interfazea da.



Irudia 4.26: Galderak zuzendu erabilpen kasuaren fluxu diagrama.

Zuzenketak gorde

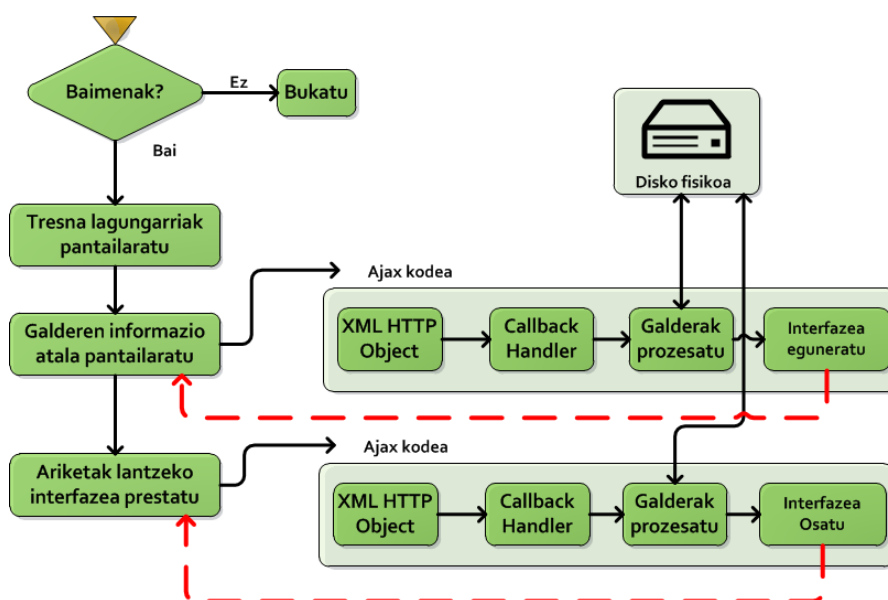
Interfazean galderak ezabatu edo zuzendu ostean aldaketak gordetzeko erabilgarri den funtzionalitatea.



Irudia 4.27: Zuzenketak gorde erabilpen kasuaren fluxu diagrama.

Ariketak landu

Aplikazioan atzigarri dauden fitxategien gainean ariketak lantzeko funtzionalitateak biltzen dituen interfazea da.



Irudia 4.28: Ariketak landu erabilpen kasuaren fluxu diagrama.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Automatikoki zuzendu

Galderen sorkuntza garaian identifikatu den emaitza -galderaren emaitza- eta erabiltzaileak erantzun gisa idatzi duen esaldiaren arteko alderaketaz arduratzen den erabilpen kasua. Esaldien alderaketa egiteko itzulpengintzan erabiltzen diren bi metrika erabili dira, eta metrika hauek itzultzen duten bat etortzearen arabera emaitza ontzat edo txartzat jotzen da. Jatorriz metrika hauek esaldien arteko urruntasuna neurtzeko erabiltzen dira, aitzitik, kasu honetan, sintagmen arteko urruntasuna neurtzeko erabili dira; hala eta guztiz ere, proiektu honetan urruntasunaren kontzeptuan saiakera txiki bat egitea baino ez da izan helburua, eta horregatik, implementatutako eredu egokia dela uste da. Esaldiak alderatzeko prozesua bi geruzatan antolatu da:

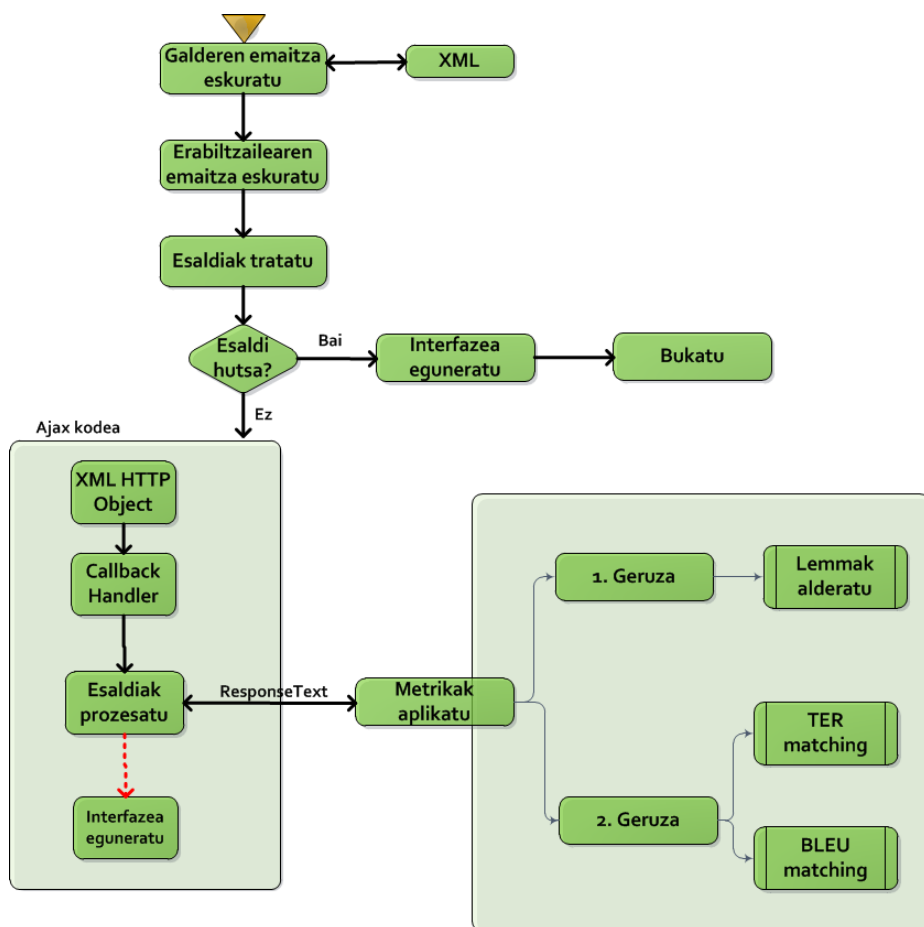
- 1. Geruza - Lemma matching

Esaldiko hitzen lemak lortzeko Morpheus zerbitzuaz baliatuz, esaldiko hitzen arteko lemen bat etortzea neurtzen da.

- 2. Geruza - BLEU eta TER (Translation Error Rate) matching

BLEU: Bi esaldien arteko antzekotasunak neurtzeko metrika, esaldien antzekotasunaren arabera 0 eta 1 artean dagoen balioa itzultzen duena.

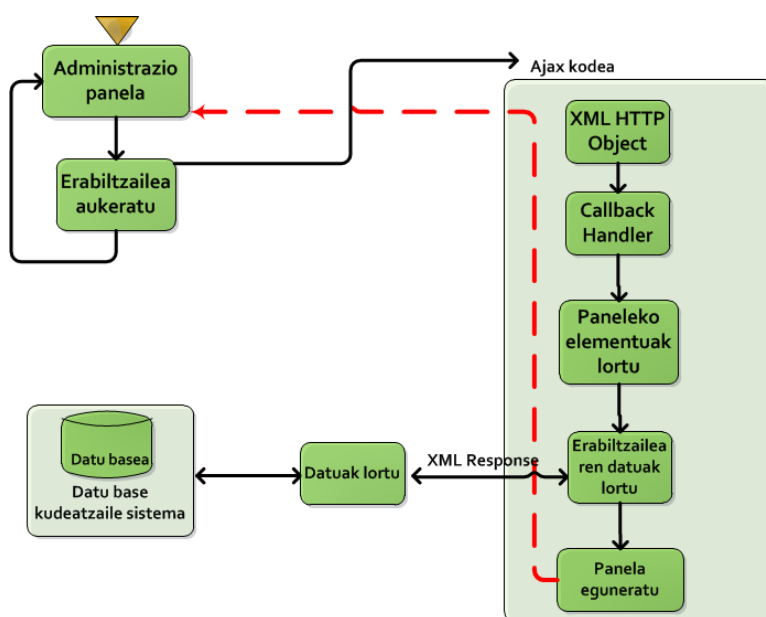
TER: Esaldiak berdinak izateko egin beharko liratekeen editatze kopurua neurtzeko erabiltzen den metrika, honek ere 0 eta 1 artean dagoen balioa itzultzen du.



Irudia 4.29: Automatikoki zuzendu erabilpen kasuaren fluxu diagrama.

Erabiltzaileen fitxategiak ikusi

Erabiltzaileak kudeatzeko funtzionalitaterik garrantzitsuenak, administratzaileari aukeratutako erabiltzaileak aplikazioan dituen fitxategi publiko eta pribatuak ikusteko aukera ematen diona. Gainera, fitxategien gainean kudeaketa eragiketarako burutzeko aukerak eskaintzen ditu, esaterako: fitxategia irakurri, fitxategiaren onarpena ikusi edo aldatu eta fitxategia ezabatu.



Irudia 4.30: Erabiltzaileen fitxategiak ikusi erabilpen kasuaren fluxu diagrama.

4.7 Segurtasuna bermatzen

Proiektu honen helburua sarean atzigarri egongo den sistema bat izanik, segurtasun analisi bat egitea beharrezkoa da. Analisi honetan, sistemak izan ditzakeen segurtasun zuloak ikuskatuko dira eta posible den neurrian konponbideak ematen saiatuko gara.

Segurtasun informatikoa informazio sistemen babesa bermatzea helburu duen diziplina da. Sistema bat guztiz babestea ezinezkoa izanik ere -egunero aurkitzen dira segurtasun zulo berriak-; arrisku ezberdinek sistema kolokan jartzeko probabilitatea minimora jaisteko ahaleginak egitea beharrezkoa da. Honetarako, hainbat estandar, protokolo, erreminta, arau eta metodoz baliatuko gara.

Segurtasun informatikoak informatikaren eremu asko bereganatzen ditu, izan ere, bere barnean hartzen ditu: softwarea edota software garapena, hardwarea, datu baseak edota hauen kudeaketa, informazioa, fitxategi sistema, ...; hau da, oro har erakundearentzat garrantzitsua izan daitekeen edozer-aktibo gisa ezagutzen dena-.

Esandako modura, aktiboak babestea da helburua. Proiektuaren zereginetan aktiboak babesteko kontuan izango ditugunak ondorengoak dira: fitxategi sistemaren babesa, gordetako pasahitzen babesa, erabiltzaileen fitxategi pribatuen babesa eta SQL injection motako erasoen aurkako babesa.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Ondoren azalduko diren segurtasun neurriak kontuan izanik, edozein sistematan bermatu behar diren segurtasun ezaugarriak beteko direla uste da.

- Konfidentzialtasuna
- Informazioaren integritatea
- Autentikazioa
- Sistemaren jarraitutasuna

4.7.1 Fitxategi sistema babesten

Fitxategi sistema babesteko backup politika bat martxan jarri behar da. Gainera, segurtasun politika zuzen eta oso baten lehen urratsa ere izaten da askotan.

Gaur egun, informatika munduan batez ere, backup terminoa maiztasun handiz entzuten den kontzeptu bat da. Backup terminoak segurtasun kopiak -edo babes kopiak- egiteari deritzo. Hau da, sistema baten informazioa toki seguru batean bikoiztu eta babestea. Segurtasun kopien helburua galeren aurrean sistema informatikoa berreskuratzeko kapazitatea izatea da.

Software enpresetan esaterako, galera arazoak oso larriak dira baliabiderik garrantzitsuena informazioa baita. Hain da garrantzitsua backup politika bat ezartzea ezen badirela kasu batzuk non segurtasun istripu batzuen ondorioz enpresak desagertu egin diren -aurreko egoera berreskuratzeko gai izan ez direlako alegia-.

Oro har, ordenagailuetako disko gogorretan gordetako informazioa era ziur eta iraunkorren dagoela pentsatzeko joera dago. Hau guztiz iruzurra da, izan ere, gertakari anitzek jar dezakete sistema informatikoa kolokan, esaterako: hausketak edota kolpeak, hardware akatsak, korrante elektrikoaren akatsak, ebasketak, suteak, sareko erasoak ...

Ezaugarri hauek direla eta, karrera bukaerako proiektu hau babesteko backup politika bat ezarri da. Segurtasun kopia hauek proiektuaren bizi ziklo osoan zehar burutuko dira.

Backup politika

Proiektuaren bizi ziklo osoan zehar martxan egongo den backup politika ondorengo da.

- Ordenagailu pertsonaletan dauden datuak babesteko backup politika:

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

- Astero kopia oso⁹ bat burutuko da.
 - * Gehienez ere lau kopia oso mantenduko dira.
- Egunero kopia inkremental¹⁰ bat burutuko da.
 - * Astea bukatzean eta kopia osoa burutzean, kopia inkrementalak ezabatu egingo dira.
- Zerbitzarietan¹¹ dauden datuak babesteko backup politika:
 - Egunero kopia inkrementalak burutzen dira.
 - Noizean behin kopia osoak burutzen dira.

4.7.2 Pasahitzak babesten

Garatu den aplikazioak erabiltzaile mota desberdinekin lan egiteko gai izan behar duenez; erabiltzaileen pasahitzak modu seguru batean gordetzeko erronkari aurre egin behar zaio. Pasahitzak gordetzea hasiera batean dirudiena baino ataza konplexuagoa da. Izan ere, jakina da erabiltzaile askok beti pasahitza berbera erabiltzen dutela aplikazio ezberdinetan. Horregatik, nahiz eta pasahitzen biltegiratze segurua gure sistemaren funtzionamentuari begira arazo kritiko bat ez izan, ziurrenik erabiltzailearentzat hala izango da.

Gauzak honela, hasiera hasieratik baztertu da pasahitzak testu hutsez biltegiratzea eta honen ordezkodiren bi soluzio analizatu dira:

1. Pasahitzak hash funtzio bat erabili ostean biltegiratzea, esaterako: md5, sha1 eta sha256.
2. Pasahitzak datuak zifratzeko algoritmo bat erabili ostean biltegiratzea, esaterako: DES, RSA eta DSA.

Analizatu diren bi aukera hauetatik bigarrena da hobetsi dena. Izan ere hash funtzioak ez dira gaur egun ziurrak pasahitzak gordetzeko¹², eta honen bestez, segurtasun adituek ez dute hash funtzio hauen erabilera gomendatzen pasahitzak biltegiratzeko.

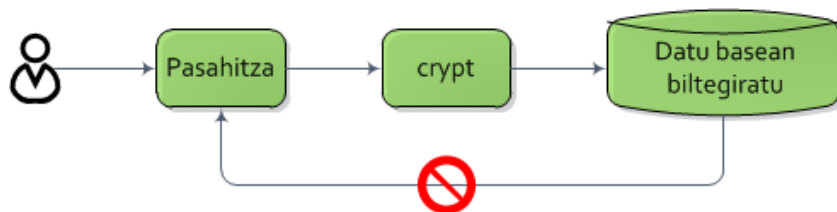
⁹Segurtasun kopia oso batek proiektu bateko informazio guztia kopiatzen du.

¹⁰Segurtasun kopia inkremental batek puntu batetik aurrerako informazio guztia kopiatzen du. Esaterako, marka hau denbora marka bat izan daiteke, adibidez: egun zehatz batetik aurrera sortutako dokumentuak.

¹¹Zerbitzarietako segurtasun kopiak CIDIR-en kudeatzen dira. Kopia hauek egiteko EMC Networker softwarea erabiltzen da.

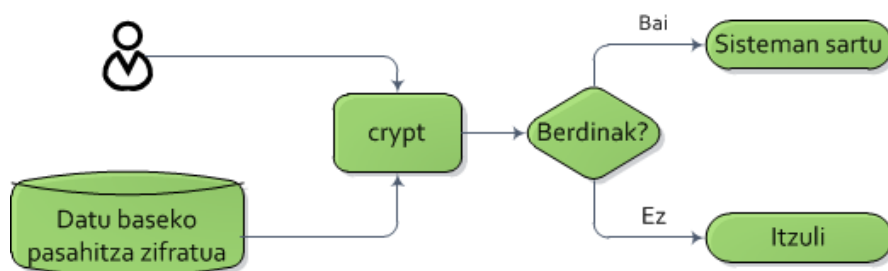
¹²Hash funtzio batek sarrera bat mapeatu eta irteera gisa hash kode bat itzultzen du. Irteera honetatik abiatuta ezin da sarrera lortu, horregatik pentsatu izan da hash funtzioak kateak era ziur batean gordetzeko baliagarriak direla. Alabaina, pentsatzeko joera hau okerra da. Gaur egun konputagailuek duten abiadura dela eta, denbora kontua da indarkeria eraso bat medio sarrera katea eskuratzea.

Hortaz, datuak zifratzeko algoritmo bat erabili ostean biltegitratuko dira pasahitzak datu basean. Zehazki PHP API-ak eskaintzen duen *crypt* funtzioa erabiliko da. Funtzio honek DES¹³ algoritmoa erabiltzen du informazioa zifratzeko. Hurrengo irudietan garatuko den login sistemaren diagramak ikus daitezke.



Irudia 4.31: Pasahitzak biltegitratzeko eskema.

Irudian ikusten den modura, pasahitz bat biltegitratu ostean ezin izango da jatorrizko katea eskuratu.



Irudia 4.32: Kautotzea burutzeko eskema.

4.7.3 Erabiltzaileen fitxategi pribatuak babesten

Proiektuaren deskribapenean azaldu den modura, aplikazioan erabiltzaile erregistratuak zein erregistratu gabeak elkarlanean arituko dira. Erabiltzaile erregistratu gabeei dagokienez, ez dute sisteman fitxategiak gordetzeko aukerarik izango. Aldiz, erabiltzaile erregistratuek fitxategiak sisteman gordetzeko aukera izan behar dute; eta gainera, fitxategien partekatze modua aldatzeko aukera ere bai. Edozein kasutan fitxategi baten partekatze moduak publikoa ala pribatua behar du izan. Fitxategi publikoak edozein erabiltzailek atzitu ahal izango ditu; fitxategi pribatuak ostera, egileak bakarrik atzitu ahal izango ditu.

¹³Informazio gehiago http://en.wikipedia.org/wiki/Data_Encryption_Standard

Sistemaren diseinu erabaki hauek berrikusi ondoren, bermatu behar diren egoerak deskribatzen dituen taula bat garatzeko prest gaude. Kontuan izan taulan partekatze modua azaltzen den zutabeak edozein erabiltzailearena izan daitekeen fitxategi bati egiten diola erreferentzia.

Erabiltzaile mota	Partekatze modua	Atzipena?
Administratzailea	Publikoa	Baimendu
Administratzailea	Pribatua	Baimendu ¹⁴
Erregistratua	Publikoa	Baimendu
Erregistratua	Pribatua	Baimendu b.s.b. egilea bada
Erregistratu gabea	Publikoa	Baimendu
Erregistratu gabea	Pribatua	Debekatu

Taula 4.6: Erabiltzaile, partekatze eta atzipen mota ezberdinak erlazionatzen dituen taula.

Hau bermatzeko, testu fitxategietara burutzen diren atzipen guztiak web aplikazioaren bitartez egin beharko dira. Baina honek arazo bat suposatzen du. Izan ere, web zerbitzaria *siuc01.si.ehu.es* makinan exekututzen da *apache* erabiltzailearen izenean eta ondorioz, *apache* erabiltzaileak fitxategi guztiak atzitzeko baimena izan behar du. Arazo hau ekidin ezean, erabiltzaile guztiak nabigatzailean path absolutua idatziz gero fitxategi guztiak atzitzeko aukera izango lukete. Hau saihesteko irtenbide bikoitza diseinatu da:

1. Erabiltzaile bakoitzarekin lotuta egongo den ausazko kode bat erabiltzea.
2. Aplikazioaren informazioa fitxategi sistema ezberdinetan banatzea; era honetan nabigatzaile bidezko atzipena galaraziz.

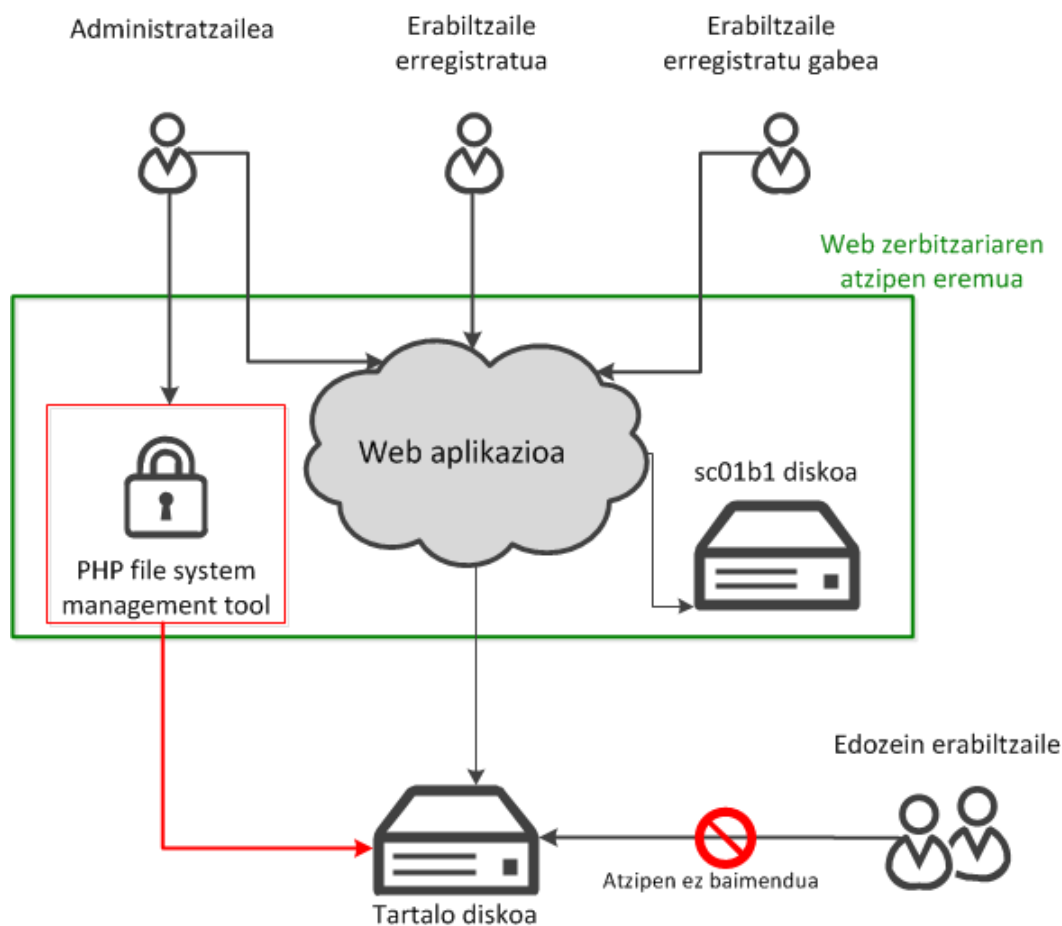
Lehenengo erabakia, segurtasun neurri gehigarri bat dena, erabiltzaile baten sistemara igo dituen fitxategietara atzitzeko zailtasun gehigarri bat izatea du helburu. Honela erabiltzaile izena ez ezik, ausazko kode hau ere beharrezkoa izango da datuak atzitu ahal izateko. Ausazko kode hau erabiltzailea erregistratzen denean sortuko da eta web aplikazioak bakarrik eskuratu ahal izango du.

Bigarren erabakiak aldiz, aukera ematen du aplikazioaren informazioa web zerbitzariaren atzipen eremutik kanpo jartzeko. Honela, *apache* erabiltzaileak ez du arazorik izango informazioa atzitzeko -zerbitzariaren aldeko

¹⁴Nahiz eta fitxategia beste erabiltzaile batena izan, atzipena baimentzen da. Honek horrela behar du izan administrazio lanak burutzeko.

programazio hizkuntza erabiliz betiere-; baina era berean, informazio hau ezin izango da nabigatzailetik atzitu. Ondorioz, testu fitxategietara egindako atzipen guztiak hala moduz web aplikazioa erabiltzera behartuta egongo dira.

Ondorengo irudian aipatutako ebazpenaren diagrama azaltzen da, azalpenak argitzen lagunduko duelakoan.



Irudia 4.33: Erabiltzaileen fitxategiak babesteko ebazpenaren diagrama.

Aitzitik, ebazpen honek ondorio kaltegarri bat dakar aplikazioaren garapenerako. Izan ere, banatutako diskoan gordetako datuak apache erabiltzailearen izen eta taldean egon behar dutenez; apache-k bakarrik -eta ez beste inork- atzitu ahal izango ditu. Garapen garaian fitxategi hauek atzitzea beharrezkoa denez, PHP bidezko fitxategi sistema kudeatzaile baten beharra nabaritzen da. Honela, web bidez fitxategi hauek kudeatu ahalko dira.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Zeregin hauetarako *PHPfm (PHP File System Management Tool)*¹⁵ kudeatzailea erabiliko da. Baina aldi berean horrelako kudeatzaile bat web bidez eskaintzea arriskutsua denez *.htaccess (hypertext access)*¹⁶ bidez babestuko da katalogoa.

4.7.4 SQL injection-en aurka babesten

SQL injection¹⁷ kode txertaketa bidez informazioa baimenik gabe erauzteko teknika bat da. Teknika hau datu baseen erabileran dauden segurtasun zuloz baliatzen da. Oso ezaguna bilakatzen ari da azkenaldian eta oso arriskutsua izan daiteke sistema eraso hauen aurka babesten ez bada. Izan ere, erasotzaile batek datu baseko informazio guztia erauz dezake, honek dakartzan ondorio guztiekin.

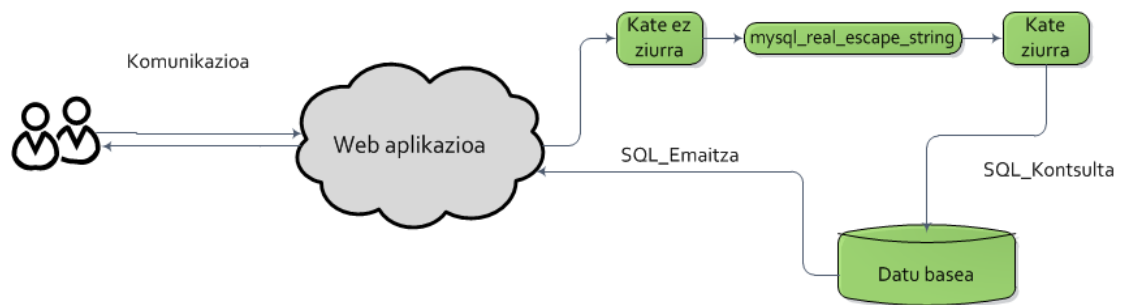
Dena den, SQL injection erasotik babestea ez da batere zaila. Programatzeko ohiturak aldatzea eta ideia bat gogoan izatea besterik ez da behar. Ideia ondorengoa da: erabiltzaileek sistemari bidaltzen dioten informazio oro ez ziurtzat hartu behar da. Honela, edozein testu katerekin lan egin aurretik testu katea bera seguru bilakatu behar da. Hau burutzeko metodo ezberdinak daude eta kasu honetan PHP API-ak eskaintzen duen metodo bat erabiltzea erabaki da.

Metodoak *mysql_real_escape_string* du izena, eta honekin, garatutako aplikazioa SQL injection motako erasoen aurka babestua egongo da. Izan ere, metodo honek arriskutsuak diren karaktere guztiak babesten ditu. Hortaz, aplikazioko datu sarrera guztiak babestea nahikoa izango da eraso mota hau ekiditeko.

¹⁵<http://phpfm.sourceforge.net/>

¹⁶<http://en.wikipedia.org/wiki/.htaccess>

¹⁷Jakin minarekin geratzen den irakurleari aipatu txosten honen eranskinetan *SQL injection-en adibide erraz bat* azpiatalean (I.b) SQL injection kasu simple bat aztertzen dela.



Irudia 4.34: SQL injection ekiditeko eskema.

5 Kapituluia

Software garapena

Gaien Aurkibidea

5.1	Softwarea garatzeko prozesua	112
5.1.1	Garapen iteratibo inkremental prototipatua	112
5.2	Web aplikazioa munduratzea	113
5.2.1	Search Engine Optimization	114

Atal honetan implementazioaren inguruko aipamen garrantzitsuenak azalduko dira, euren artean: zein software prozesu erabili den sistema garatzeko eta nola munduratu den aplikazioa.

5.1 Softwarea garatzeko prozesua

Softwarea garatzeko balio duen metodoari software prozesua deitzen zaio. Software prozesuak uneoro garapena bideratuko eta jarraitu behar duen jardueren sekuentzia adieraziko du. Argi dago software prozesuaren helburua produktuak garatzea dela, eta honetarako, urrats desberdinak ematea beharrezkoa izango da produktuaren kostua eta kalitatea egokia izatekotan. Helburu hauek lortzeko desberdin baina, aldi berean, osagarri diren azpi prozesuak erabili ohi dira: garapen prozesua, gauzatu behar diren atazak zehazten dituen; kudeaketa prozesua, jardueren planifikazioa kontrolatzen duena eta konfigurazio prozesua, aldaketan eta produktuaren integritatea kontrolatzen dituen. Dena den, jarduera sekuentzia nahiko abstraktua izanik, proiektu bakoitzak bere instantziazioa eman behar dio jarduera generiko bakoitzari, proiektuaren bizi zikloko nondik norakoak zehatz azalduz.

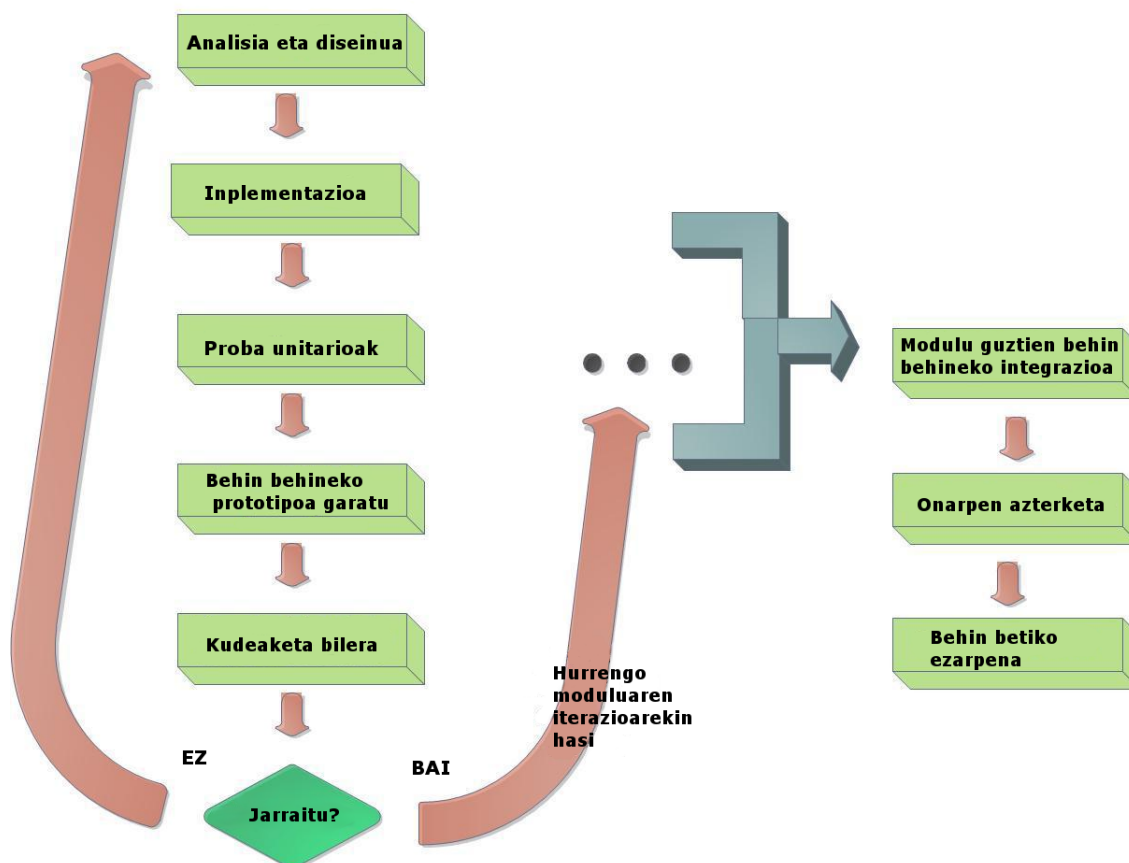
5.1.1 Garapen iteratibo inkremental prototipatua

Proiektua garatzeko *metodologia arina* erabiliko da; honela: tarte labur batean software funtzionala garatzeko eta aldaketak bukaerako produktuari eraginkortasunez egokitzeko aukera izango baita. Metodologia hau egikaritzeko kideen arteko elkarrekintza oinarritzakoa da; honetarako, proiektuaren deskribapenean zehaztu den lan metodologia bizi ziklo osoan zehar mantentzea funtsezkoa izango da.

Jarraitzen den software prozesua *Extreme Programming*¹ metodologiaren egokitzapena da. Era honetan, garapen inkrementalean oinarrituta, aplikazioa moduluetan zatitzen da eta modulu bakoitza miniaplikazio baten modura garatzen da. Hauetako miniaplikazio bakoitzak bere betekizun, diseinu eta implementazio partikularrak izango ditu. Ondorioz, iterazio bakoitzaren amaieran modulu hauetako bakoitzak bukaerako aplikazioaren funtzionalitateak areagotuko ditu. Gainera, garapena prototipatua izateak bezeroen betekizunak hasiera hasieratik eskuratzeko aukera ematen du; era honetan, nahiz eta sistema guztiz funtzionala ez izan bizi zikloan zehar aurkitutako akatsekin sistema birdiseinatuko da.

Ondorengo irudian iterazio bakoitzean burutzen diren atazak zerrendatzen dira:

¹http://en.wikipedia.org/wiki/Extreme_programming



Irudia 5.1: Softwarea garatzeko iterazioen eskema.

5.2 Web aplikazioa munduratzea

Behin web aplikazioa garatu ondoren erabiltzaileak webgunean lanean aritzea da gure helburu nagusia. Baina, argi dago zerbait egin beharrez gaudela tresna erabiltzaileengan gerturatzeko. Gaur egun, internetek eskaintzen duen sarearen hedadura kontuan izanik, besterik egin ezean, nekez lortuko genuke erabiltzailearen bat edo beste gure webgunean aritzea. Horregatik, Search Engine Optimization estrategia martxan jartzea erabaki dugu garatutako aplikazioa munduratzeko eta erabiltzaileak erakartzeko helburuarekin.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

5.2.1 Search Engine Optimization

Search Engine Optimization (SEO) interneteko bilatzaileek eskaintzen dituzten bilaketan emaitzetan kokaleku ona lortzeko ahalegina da. Oro har, zenbat eta hobekiago kokatu kontsulten zerrenda hauetan, orduan eta baliagarriagoa da webgunearentzat; bilatzailearen erabiltzaileak maizago bisitatuko baitute ongi kokaturik dagoen webgunea.

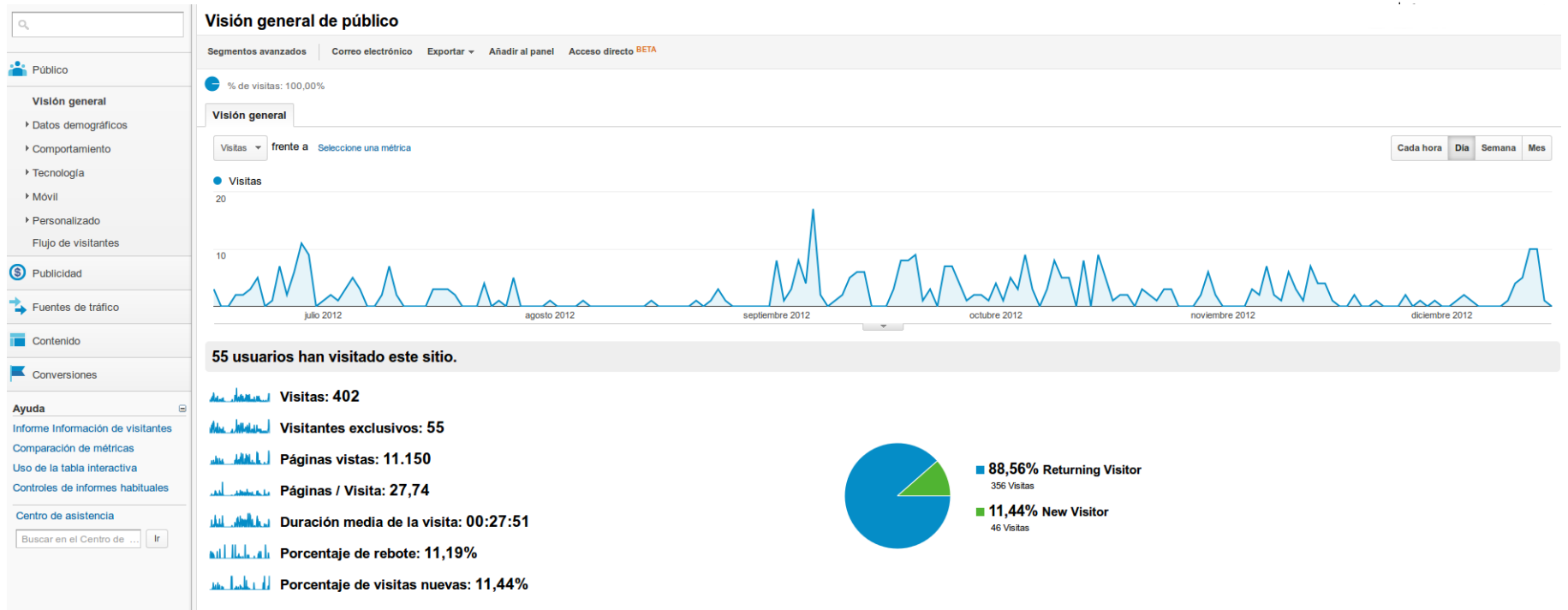
Webgunearen erabilera areagotzeko helburuarekin, ondorengo bilatzaileetan indexatu dugu webgunea

1. Google
2. Yahoo
3. MSN
4. Open Directory Project²

Gainera, bilatzaile hauek web garatzaileei oso baliagarriak zaizkien tresnak eskaintzen dizkiete webguneen kudeaketari begira, esaterako: Google Analytics eta Bing web administratzailea. Aplikazio hauen bitartez webguneen kudeaketa modu automatizatu eta erraz batean egin daiteke. Eskaintzen dituzten funtzionalitate guztien artean hauek dira aipagarrienak:

- Webgunean sor daitezkeen arazoen inguruko jakinarazpen zerbitzu automatikoa.
- SEO estatistiken laburbilduak.
- Webguneak jasotzen duen trafikorearen inguruko estatistikak.
 - Datu demografikoak.
 - Bisitaldien maiztasuna eta interakzio informazioa.
 - Erabilitako nabigatzailearen, sistema eragilearen eta sarearen inguruko informazioa.
- Webgunea optimizatzeko aholkuak, hala nola: izenburu egokiak idazteko laguntzak, hitz gakoaren inguruko informazioa, HTML atributuen erabilera gomendioak, ...

²<http://www.dmoz.org/>



Irudia 5.2: Google Analytics web kudeatzailea.

6 Kapituluia

Ebaluazioa, emaitzak eta arazoak

Gaien Aurkibidea

6.1	Galdera sortzaile sistemaren ebaluazioa	118
6.1.1	Emaitzak	118
6.1.2	Errore analisia	120
6.1.3	Ondorioak	120
6.2	Web interfazearen ebaluazioa	120
6.2.1	Ebaluaketaren nondik norakoak	122
6.2.2	Erabiltzaileen iritzia	122
6.2.3	Ondorioak	125
6.3	Erabiltzaileen feedback-a	125
6.4	Proiektuan zehar izandako arazoak	127

Atal honetan proiektuan zeharreko arazoak eta ebaluazio ezberdinetan lorturiko emaitzen berri emango da; baita hauetatik abiatuta lorturiko ondorioak ere.

6.1 Galdera sortzaile sistemaren ebaluazioa

Lehenik eta behin aipatu behar da ebaluazioa burutzeko IXA taldekideen parte hartzea funtsezkoa izan dela. Haien laguntzarekin ebaluazioa eskuz egiteko aukera izan baita.

Ebaluatzeko espezializaturiko korpus bat erabili da. Korpus hau euskara ikasteko helburuarekin idatziriko fitxategi bilduma bat da, eta -CEFR¹-ek esandakoarekin bat- zailtasun maila desberdinetan sailkatuta dago. Fitxategi bilduma hauen guztien artean erdi mailakoak erabiltzea erabaki zen galderak sortzeko oinarri gisa. Oro har, dokumentuen zailtasun maila zehaztea ataza erraza ez bada ere, kasu honetan, korpusa irakasle adituen eskuzko lanen bitartez sailkatua izan da.

Esperimentu honetan, jatorrizko fitxategiarekin erlazionaturiko hainbat galdera sortu eta ondoren emaitzak aztertu ziren². Azterketan sorturiko galderen zuzentasun sintaktikoari eta gramatikalari arreta jarri zitzaion, era honetan, sistemaren galderak sortzeko gaitasuna ebaluatuz. Honetaz gain, adituak erabilitako galdetzaileen zuzentasuna eta galderek jatorrizko testuarekiko zuten interesa ere aztertu zuen. Azkenik, ebaluatu ostean, emaitzak aztertu eta ondorioak zein etorkizuneko lanak zehaztu ziren.

Atal honetan argitu behar da ebaluaketa garapenean zehar egin dela, eta, honen ostean, galdera sortzaile (QG) sisteman hobekuntza gehiago diseinatu eta inplementatu direla. Ondorioz, QG sistemaren bigarren bertsioan aipatu diren hobekuntza guztiak ez dira ebaluaketa honetan islatzen.

6.1.1 Emaitzak

Behin dokumentuak prest izanik, galderak masiboki sortu eta hauen guztien artean 180 ausaz aukeratu ziren. Zehazki ebaluatu behar zen deklinabide kasu bakoitzeko hogeita hamar galdera. Landuriko deklinabide kasuak ondokoak dira: ablatiboa, absolutiboa, adlatiboa, adlatibo bukatuzkoa, datiboa, destinatiboa, ergatiboa, inesiboa eta sozietiboa. Azpian, emaitzetatik abiatuz lorturiko datuak ikus daitezke:

¹Common European Framework of Reference for Languages

²Eranskinetan azaltzen den *Galdera sortzaile sistemaren ebaluaketako emaitzak* atalean (I.c) ebaluazio honen emaitza batzuk azaltzen dira, aipatzen diren kontzeptuak ulertzeko lagungarri suertatuko delakoan

- Hautagaien aukeraketa heuristikoari dagokionez: hautaturiko erantzunen %10.61a entitate izenduna da (%3.31a pertsona, %4.6a tokia eta %2.7a erakundea), %7a biziduna eta %56.05a bizigabea da, %8.00a roll semantikoetatik abiatutakoa da eta gainerako %18.33a informazio morfologikoan soilik oinarritutakoa.
- Galderen zuzentasun gramatikalari dagokionez: %39.08a galdera guztiz zuzenak dira, inolako zuzenketarik behar ez dutenak; %6.9a gramatikalki zuzena da, baina ohiz kanpokoa; %13.22a arazo gramatikalen bat duten galderak dira eta %40.80a baztertua izan da arazo gramatikal larriak direla eta. Honetaz gain, aipatu behar da galdera zuzenen portzentaia desberdina dela heuristiko bakoitzaren kasuan; esaterako, bizidun eta leku entitateekin portzentaia hobeak lortu dira, %55.56 eta %42.86 hurrenez hurren.
- Deklinabide kasuei dagokienez: ergatiboa eta adlatiboa izan dira zailtzarrik gabe emaitza okerrenak izan dituztenak; ergatiboaren kasuan sortutako galdera guztien %70a baztertuak izan dira eta adlatiboaren kasuan %63.15a. Aitzitik, absolutibo, adlatibo bukatuzko, destinatibo eta inesiboaren kasuan sortutako galdera gramatikal zuzenen portzentaia %50a baino handiagoa da.
- Soilik galdetzaileen zuzentasunari dagokionez, galdera guztiei esleituri-ko galdetzaileen %63.21 zuzena da. Gainera, oso nabaria da portzentaia hau asko igo edota jaisten dela hautagaia aukeratzeko erabili den heuristikoa dela medio. Esaterako, semantika hiztegian oinarritutako galdetzaileen esleipen zuzena %88.24koa eta %80.68koa da bizidun eta bizigabeentzat hurrenez hurren. Hautagaia informazio morfologikotik abiatuta aukeratu bada ordea, galdetzaileen esleipen zuzenaren portzentaia %26.83ra jaisten da.
- Galderen interesari dagokionez, jatorrizko testuarekiko lotura garrantzitsua egiten duen galderen portzentaia %74.39koa da. Gainera, aipatu beharrekoa da semantika hiztegian zein entitate izenetan oinarrituriko galderetan, kasu guztietan galdera interesgarriak izan direla; eta aldiz, informazio morfologikotik abiatutako galderetan %60.98a bakarrik da interesgarria. Kasu honetan ere, adlatiboa da emaitza okerrenak bereganatu dituen deklinabide kasua (%52.63).

6.1.2 Errore analisia

Emaitzen eskuzko analisiak zein ebaluazioa burutu duen adituarekin egikaritu diren bilerek, errore analisi bat aurrera eramateko aukera eman digu.

Besteak beste, galderen zuzentasun gramatikala ebaluatzerakoan arazo desberdinak dituzten egoerak identifikatu dira:

1. Gramatikalki zuzenak diren galderak baina hiztun batek sekula erabiliko ez lituzkeenak.
2. Akats ortografikoak dituzten galderak.
3. Morfologiari dagokionez gaizki eratuta dauden galderak.
4. Idatzizko euskararekin ez baizik eta ahozkoarekin lotzen diren galderak.
5. Puntuazio ikurrekin arazoak dituzten galderak.
6. Hitzen ordena desegokia duten galderak.

6.1.3 Ondorioak

Testu fitxategi batetik abiatuta euskaraz galderak sortzeko gai den sistema baten hastapena garatu da; semantika hiztegi, izendun entitate, roll semantiko eta morfologian oinarrituta esaldi mailan galderak egiten saiatzen dena.

Ebaluazioko emaitzetatik jaso dugun informazio garrantzitsuena galderen %39.08a zuzena dela izan da, gainera, portzentai hau handitzen da -edo txikitzen- erabili den heuristikoen arabera.

Argi dago etorkizuneko lana heuristikoen hobekuntzari lotuta bideratu behar dela. Gaur egun, galderak sortzeko sistemak bizidun eta bizigabeei preferentzia ematen die, eta honek horrela behar du izan, kasu hauetan lortu baitira emaitzarik onenak. Hala ere, beste heuristikoei dagokienez, esaterako roll semantikoak automatikoki eskuratzeko gai izango bagina, lortutako emaitzak asko hobetuko lirakeela uste da. Ondorioz, helburua galderen sorkuntza hobetuko duten NLP tresna berrien analisi eta integrazioan arreta jartzea da. Jakina da zenbat eta linguistikoki aberatsagoa den analisia izan, orduan eta galdera hobeak sortuko direla; horregatik, izendun entitate zein roll semantikoen arloan aurrerapenak egitea oinarritzkoa da.

6.2 Web interfazearen ebaluazioa

Seneko aplikazioa ingurune erreal batean praktikan jartzeko, ebaluatzeko eta ondorioak biltzeko asmoz, eskola baten parte hartzea eskatu da. Ebaluazio

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

honetan aplikazioaren erabilera baloratu nahi denez, ez gara sortutako galderen zuzentasun gramatikalean sartuko; aitzitik, webgunearen interfazea, erabilgarritasuna eta eskaintzen dituen aukerak ebaluatu nahi ditugu.

Aplikazioaren ebaluaketa Martuneneko Arantzazuko Ama³ eskolan egin da. Lehen Hezkuntzako bigarren zikloko (LH4 maila) hamahiru neska-mutilez osatutako ikastalde batek parte hartu du. Ebaluaketa informatika ikasgelan egin da, eta ordu eta erdiko iraupena izan du guztira. Bertan, egoera bat aurkeztu -eskuliburu baten laguntzaz betiere- eta haurrak web aplikazioan lanean aritu dira azaldutako pausoak eginez. Saioan zehar ebaluatzaileek -proiektuko zuzendariak eta ikasleak- planteatutako ariketak landu dituzte eta azkenik galdeketa txiki bat erantzun dute. Galdeketa honek erabiltzaileen iritziak eta aplikazioaren ingurukoak jasotzea du helburu. Jarraian, ebaluaketa saioaren planifikazioa azaltzen da:

Planifikazioa

- Ebaluaketa taldea bederatzi eta hamar urte bitarteko haurrek eta ebaluatzaileek osatutako dute.
- Ebaluatzaileen partaidetza ezinbestekoa izango da dinamika aurrera eraman ahal izateko. Hauek izango baitira material guztia prestatzeko eta ebaluazioan sortzen diren galdera eta arazo guztiak konpontzeko arduradunak.
- Ebaluaketa saioa prestatzeko erabili den materiala ondorengoa da: haurrentzat egindako eskuliburua, azalpen argiz eta eman behar duten pauso bakoitzeko hainbat irudiz osatua; haurrentzako bost galdera motzez osatutako galdeketa eta irakasleentzako galdeketa, zazpi galderaz osatua.
- Ikasleak binakako taldetan banatu, eskuliburuak eman eta hauetan agertzen diren pausoak jarraitzeko eskatuko zaie. Lanak binaka egiteko metodologia hezkuntza ziklo hauetan erabiltzen den teknika izanik, web aplikazioaren ebaluaketa egiteko ere aproposa izan daitekeela uste da; era honetan, taldeko ikasle aurreratuenak tutore lanak egingo ditu bere ikasketa kideari lagunduz, ondorioz, planteatzen diren ariketez gain, umeen arteko elkarlana ere sendotzea espero da.
- Saioan egin beharrekoa ondorengoa izango da: lehendabizi, ebaluatzaileek egingo duten azalpena arretaz entzun. Ondoren, ebaluatzaileek prestatutako fitxategi publiko baten galderak erantzun. Honetaz

³<http://www.arantzazukoama.org/>

gain, eta denborarik bada, ariketa aurreratu bat egiten saiatu behar dira haurrak: sisteman kautotu eta fitxategi pribatu bati galderen bat edo beste gehitu, gainera, fitxategiaren partekatzea aldatu beharko dute, haien ikaskideei beraiek sortutako galderak erantzuteko aukera emateko. Azkenik, aplikazioa nahi haina “saltxeatzeko” eskatuko zaie, halaber, txatean ibiltzeko proposatuko zaie; eta bukatzeko, iritzigaldeketa emaitzak jasoko dira.

6.2.1 Ebaluaketaren nondik norakoak

Jarraian, eskolako hurrek laborategi saioan zehar webgunean egindako lanak laburbiltzen dira:

1. Ikasle guztiek ipuin baten gainean proposatutako lau galdera erantzun dituzte -galdera sortzaile sistemak sortutakoak-.
2. Ikasle guztiek, sisteman kautotuz, galdera bat edo bi gehitu dizkiote hasieran proposatutako ipuinari.
3. Ikasle guztiek beste ikaskideren batek eskuz sortutako galdera berriak erantzun dituzte.

6.2.2 Erabiltzaileen iritzia

Ondorengo lerroetan ebaluaketan egindako galdeketa emaitzak laburbiltzen dira. Lehendabizi, hurrek galdeketa erantzundako emaitzak aztertuko ditugu, eta, ondoren, laborategiko saioan gurekin egon diren irakasleek erantzundakoa. Haurren galdeketa bost galdera motzez osatua dago. Hona galderak:

1. Gustatu al zaizu aplikazioa?
 - a) Bai, asko
 - b) Ez gehiegi
 - c) Ezer ez Zergatik?
2. Zaila iruditu al zaizu erabiltzeko modua?
 - a) Ez, segituan ulertu dut
 - b) Hasieran pixka bat kostatu zait ulertzea
 - c) Ez dakit nola erabiltzen den

3. Sortutako galderak testua ulertzen lagundu al dizute?
 - a) Bai
 - b) Pixka bat
 - c) Ez
4. Gustatuko zitzaizun irakasleek metodo hau erabiltzea testuak lantzeko gelan?
 - a) Bai, askoz atseginagoa da
 - b) Ez, nahiago dut betiko metodoa (testua eta galderak paperean izatea)
5. Nola baloratuko zenuke aplikazioa 1etik 10era? Zergatik?

Haurrek galdeketa honi erantzundakoa ondorengoa da. Proposatutako lehen galderari, haur guztiek lehenengo aukera hautatu dute, aplikazioa asko gustatu zaiela alegia. Aukera hau hautatzeko arrazoiak, aldiz, desberdinak dira: paperik eta libururik izan gabe irakurtzeko eta lan egiteko aukera ematen duelako, etxetik ere lan egiteko aukera ematen duelako, dibertigarria baina aldi berean ikastera bultzatzen zaituelako, oso ondo aritu direlako laborategi saioan, landutako ipuina oso polita zelako, ... Proposatutako bigarren galderari dagokionez, hamahiru haurretatik zazpik hasieran aplikazioa ulertzea pixka bat kostatu zaiela onartu dute; gainerakoak, ordea, hasieratik segituan ulertu dutela diote. Honek, etorkizunari begira nabigazio eta ulergarritasuna hobetzeko ahaleginak egitera behartzen gaitu. Hirugarren galderari dagokionez, zortzi haurrek proposatutako galderak testua ulertzeko lagungarriak izan direla aipatu dute, gainerakoek, aldiz, pixka bat lagungarriak suertatu zaizkiela. Laugarren galderari dagokionez, haur guztiek testuak gelan lantzeko sistema hau nahiago dutela erantzun dute; testuak eta galderak paperean izatearen kontra. Azkeneko galderari dagokionez, zortzi haurrek web aplikazioa hamar batekin baloratu dute, beraien iritziz oso ondo baitago; beste hiru haurrek, bederatzi batekin baloratu dute aplikazioa, beraien iritziz txatak pribatua izan behar lukeelako, eta azkenik, beste bi haurrek, baloraziorik eman gabe, aplikazioa oso ongi dagoela eta oso ondo pasa dutela diote.

Oro har, begi onez ikusi dute aplikazioa eta gustukoa zutela adierazi digute laborategi saioan zehar ere. Nire aldetik ere, eskerrak eman behar dizkiet guztiei, eta lanean bikain aritu direla azpimarratu nahiko nuke.

Emaitza hauek ikusi ondoren, laborategiko saioan parte hartu duten irakasleen galdeketa aztertuko dugu, hona galderak:

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

1. Erabilgarria ikusten al duzu aplikazioa?
2. Zer zailtasun antzeman dituzu webgunean?
3. Zer aldatuko/hobetuko zenuke? Tresna lagungarrien atalean funtzionalitate berriren bat edo beste gehituko zenuke?
4. Zer da gehien gustatu zaizuna?
5. Zer moduz ikusi dituzu haurrak lanean?
6. Tokia ikusten al diozu eskola batean?
7. Nola baloratuko zenuke aplikazioa (1etik-10era)? Zergatik?

Lehenengo galderari dagokionez, saioan gurekin lanean aritu diren bi irakasleek aplikazioa erabilgarri ikusten dutela diote, eta haurren erritmo pertsonalei behar bezala erantzuteko aukerak ematen dituela azpimarratu dute. Bigarren galderari dagokionez, weban aurkitu dituzten zailtasunak hauek izan dira: aukera gehiegi atzigarri izateak erabilera nahasten duela eta koloreen bidezko laguntzak faltan bota dituztela, honetaz gain, aplikazioaren akats analisi sakona egiteko denbora gehiago beharko luketela aipatu digute. Hirugarren galderari dagokionez, bata, honi behar bezala erantzuteko denbora gehiago beharko duela dio, eta besteak, aldiz, ondorengo puntuak zerrendatu ditu: galdera guztiak hasieratik aurkeztu beharrean, hobe litza-tekeela galdera batean zentratzea, eta, horrekin amaitzean, hurrengora salto egitea; galderak oker erantzutean automatikoki laguntzaren atala koloretzatzea; aukera batzuk kentzea nahasmen gutxiago eragiteko eta interfazea umeentzako pixka bat erakargarriagoa bilakatzea, izan ere, filtroaren atala-deklinabide kasuak ikusten diren atala- gehiegizkoa dela iruditu zaie. Jarraitzeko, laugarren galderari dagokionez, gehien gustatu zaiena ondorengoa izan da: galdera berriak sortzeko aukera eta materiala kudeatzeko baliabideak oso interesgarriak direla eta, honetaz gain, irakasle bakoitzak bere material autonomoa erabiltzeko aukera ona dela azpimarratu dute. Bosgarren galderari dagokionez, biek haurrak lanean gustura eta jarrera aktiboaz ikusi dituztela diote. Seigarren galderari dagokionez, eskoletan ageri diren tresna informatikoen beharrak azaldu dizkigute, eta, gainera, lan egiteko erritmo desberdinak lantzeko aplikazio aproposa dela adierazi dute. Azkeneko galderari dagokionez, irakasle batek, oro har, ideia ona dela dio, baina guztiz funtzionala izateko oraindik interfazea gehiago landu behar dela azpimarratu du, hau guztia kontuan izanik, zazpiko balorazioa eman dio aplikazioari. Beste alde batetik, galdeketa erantzun duen beste irakasleak, aplikazioa interesgarria izan arren, umeentzat erabilgarria izateko interfazeko atal batzuk

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

aldatuko zituela aipatu du, eta aplikazioa zortziko puntuazioarekin baloratu du.

6.2.3 Ondorioak

Ebaluaketak hasiera batean pentsatzen genuena baino denbora gehiago⁴ behar izan duen arren, oso positiboa izan dela uste dut. Bai ebaluaketa prestatzeko egin diren bileretan aurkitu diren arazo eta identifikatu diren behar berriengatik, baita garatutako aplikazioa ingurune errealista batean martxan jarri delako, eta honekin batera, behar errealak identifikatu⁵ direlako.

Honetaz gain, eskolako irakasleen⁶ parte hartzea eta kolaborazioa erraldoia izan da, eta, oro har, proiektua atsegina eta erabilgarria gertatu zaie. Alabaina, aplikazioa guztiz funtzionala izateko oraindik aldaketa asko egin behar badira ere, eskoletan tokia ikusten diote mota honetako aplikazio informatikoei, lanak automatizatzeko eta ikasleen zereginak neurrira egokitzeko aukera ematen dutelako.

Martuteneko Arantzazuko Ama eskolan saio bakarra egin den arren, esperientzia positiboa izan da eta lanean jarraitzeko zein urrats berriak emateko aukera paregabea izan da.

6.3 Erabiltzaileen feedback-a

Atal honetan unibertsitateko boluntarioekin egindako *stress-test*ean sortutako arazoak eta proposamenak biltegitatu ditugu. Saio honetan hamar boluntariok parte hartu zuten, eta helburua webgunearen egonkortasuna neurtzea bazen ere, honetaz gain, webgunea eta, oro har, aplikazioa hobetzeko ideia asko proposatu ziren. Hona saioaren amaieran bildutako proposamenak:

- Webgunearen nabigazioa hobetu behar da.

Hasierara itzultzeko botoia tamainaz handitu.

Txatetik ariketak lantzeko txokora zein fitxategiak kudeatzeko txokora zuzenean joateko estekak gehitu.

⁴Proiektuaren birplanifikazioaren atalean (7. kapitulua) estimatutako eta benetan behar izan diren denbora tartek aztertu daitezke

⁵Etorkizuneko lanetan (8.2 azpiatala) gerora begira inplementatu beharreko zeregin guztiak zerrendatzen dira

⁶Lerro hauetan eskerrak eman nahi dizkiet Maleni eta Felixiri eskaini diguten denbora eta arreta guztiagatik

Webaren goialdean aurkezten diren bi irudiak esteka bihurtzea. Era honetan, IXA taldearen irudiak IXA taldeko webgunera joateko aukera emango luke, eta, Senekorenak aldiz, webgunearen hasierara itzultzeko.

Fitxategiak kudeatzeko txokoan eta ariketak lantzeko txokoan azaltzen diren estekak behealdean ere kopiatzea, galdera asko egonik, goraino igotzea astuna egiten baita.

Edozein unetan bi txokoetara joateko estekak agerian izatea.

- Web interfazea aldatzeko proposamenak.

Tresna lagungarrien leihoa tamainaz txikitu edo tokiz aldatu behar da, izan ere, pantaila txikietan lan eremuaren gainean pantailarazten da.

Fitxategiaren prozesuak bereganatuko duen denbora tartea aurreikusten saiatzea, eta, era honetan, erabiltzaileari itxaron behar duen denbora zehaztasun handiagoz ematea.

Txatean "scroll" bat gehitu behar da.

Txat erabilpena leiho berri batean irekitzea.

Tresna lagungarrien artean agertzen den azken erabilpena, testua leiho berrian irakurtzekoa alegia, beste toki batean pantailaraztea.

Testua igotzeko atalean, gaia eremua ez dago argi. Hobe litzateke eremu hau bi azpieremutan banatzea: alde batetik, testu mota -antzerkia, berriak, ipuina, ...-, eta bestetik, gaia -erlijioa, ekonomia, fisika, ...-. Honetaz gain, idazlana aukera ere sartu beharko litzateke, izan ere, aplikazio hau ikasleek ere erabiliko badute, interesgarria litzateke beraiek idatzitako testuak lantzeko aukera izatea.

Zuzen eta oker erantzun diren galdera kopurua zenbatzea.

- Aplikazioari begira proposatutako hobekuntzak eta arazoak.

Aplikazioa erabiltzeko informazio gehiago gehitzea.

Ariketan erantzuteko atalean, "Erantzun ondorengo galderei. Erabili itzazu testuan agertzen diren hitzak." edo tankera honetako azalpen esaldiak erabiltzea.

Internet explorer nabigatzailearekiko bateratzea ez da ona.

Automatikoki sortutako eta eskuz sortutako galderek nabarmentzea eta tratamendu desberdina ematea.

- Galdera sortzailearen inguruko arazoak.
 - Puntuazio ikurren inguruko tratamendua ez da ona.
 - Galdetzaileek ez dute galdetzailearen numeroa kontuan hartzen - nor, nortzuk, ... -.
 - Elkarrizketei tratamendu berezia ematea.
 - Koordinazio motako galderei tratamendu berezia ematea.
- Emaitzak zuzentzeko atalari dagozkion hobekuntzak.
 - Zuzentzailea gehiago zabaltzea, emaitza gehiago ontzat emateko.

6.4 Proiektuan zehar izandako arazoak

Proiektu guztietan gertatzen den modura, web aplikazioaren bizi zikloan ere hainbat arazori aurre hartu behar izan diegu. Hala eta guztiz ere, ez dira ez asko ez eta larriak izan, eta, ondorioz, ez dute egindako estimazioetan berebiziko eraginik izan.

- Zerbitzarien arteko komunikazioa egikaritzeko esperotako baino arazo gehiago agertu dira, hasiera batean PHP bidez egitea diseinatu bazen ere, azkenean, Perl lengoaiatz inplementatu da.
- Web aplikazioak fitxategi sistema atzitu behar duenez, segurtasun neurri batzuk diseinatu behar izan dira. Diseinu hau hasiera batean pentsatutakoa baino konplexuagoa izan da.
- Webgunearen interfazeari dagokionez, oraindik ere nabigatzaile guztiek HTML5 eta CSS3 lengoaiak guztiz inplementatzen ez dituztenez, edo inplementazioak desberdinak direnez, nabigatzaile desberdinetan ikusten den emaitza ez da guztiz berdina.
- Proiektua \LaTeX bidez garatzeko erabakiak txostenaren idazketa atzeratu du.
- Formakuntzarako nahi adina denbora egon ez denez, honek proiektuaren inplementazioaren atal batzuetan mantso joatea izan du ondorio gisa, izan ere, web teknologiak ezezagunak nituen proiektua hasi aurretik.
- Galdera sortzailea ez ezik, ArikIturri sistema integratzeko ahaleginak egin badira ere, azkenean ezin izan da sistema webgunean integratu, pauso gehiegi ematea eta jende asko inplikatzeko eskatzen baitzuen.

- Testuak analizatzeko erabiltzen den pipet zerbitzuan esperotakoak baino arazo gehiago egon dira. Fitxategi luzeekin prozesua asko luzatzen baita -oraindik eta gehiago eskaera konkurrenteak egiten badira-, eta ondorioz, prozesurako lehenetsita dauden denbora mugak gainditu eta akatsak agertzen dira. Hau konpontzeko -arazoa aurkitzeko gehienbat- Perl LWP moduluan esperotakoa baino gehiago sakondu behar izan da, honek dakartzan denbora atzerapenarekin.
- "Demo effect" gisa ezagutzen diren arazoak ere izan genituen, ebaluaketa prestatzeko bileran bertan hain zuzen ere, izan ere, aldagai baten luzera zela eta parametro baten igorketak huts egin zuen. Arazo hau konpontzeko parametro horren bidalketa metodoa aldatzea beharrezkoa izan da.
- Unibertsitatean software librearen erabilera oso hedatuta egon arren, oro har, ez da horrela izaten. Ebaluaketa egin dugun eskolan esaterako, ez zen linux sistemarik erabiltzen. Honek webgunea erabiltzeko arazorik ekartzen ez badu ere, arazoak saihesteko fitxategiak igotzeko prozesuan ataza bat gehitzea beharrezkoa izan da. Era honetan, fitxategiak prozesatu aurretik *dos2unix* tresnarekin aurreprozesatzen dira formatu arazoak ekiditeko.
- Defektuz Emacs editoreak \sim karaktereaz amaitzen diren fitxategiak sortzen ditu backup gisa. Honek segurtasun zulo bat ekar lezake, web zerbitzariak fitxategi hauek interpretatzen ez dituelako, eta ondorioz, zerbitzariaren aldeko kodea erabiltzaileak ikusteko aukera ematen duelako. Horregatik, Emacs editorearen lehenetsitako jarrera aldatzea beharrezkoa izan da.
- Fitxategiak igotzeko HTML input etiketari value atributua aldatzea ezinezkoa da, segurtasunagatik atributu batzuen erabilera mugaturik baititu, horregatik, nabigatzailearen hizkuntzaren arabera ezartzen dio testuaren hizkuntza.
- Web zerbitzariak segurtasunagatik ezarrita duen konfigurazioaren arabera, erroreak ezkutatu egiten ditu, honek araztatze lanak zaildu ditu; arazoez jabetzeko `/var/log/httpd/error_log` fitxategia atzitu behar izan baita.
- Webgunearen atal batzuetan kodifikazio motarekin arazoak egon dira. Oro har, UTF-8 kodifikazio ereduak erabiltzen bada ere, euskaraz, ISO-8859-1 kodifikazio ereduko karaktere batzuk ere erabiltzen ditugu. Horregatik, erabiltzaile izenen kontrola egitean ustekabeko emaitzak

lortzen genituen. Arazo hau PHP funtzioak erabiliz saihestu dugu, zehazki *utf8_encode* eta *utf8_decode* funtzioekin.

- Ebaluaketari begira burutu diren azken uneko aldaketak txostenaren idazketa eta aurkezpenaren prestaketa atzeratu dute implementazio lanak egitera behartu baikaituzte.

7 Kapituluia

Proiektuaren birplanifikazioa

Gaien Aurkibidea

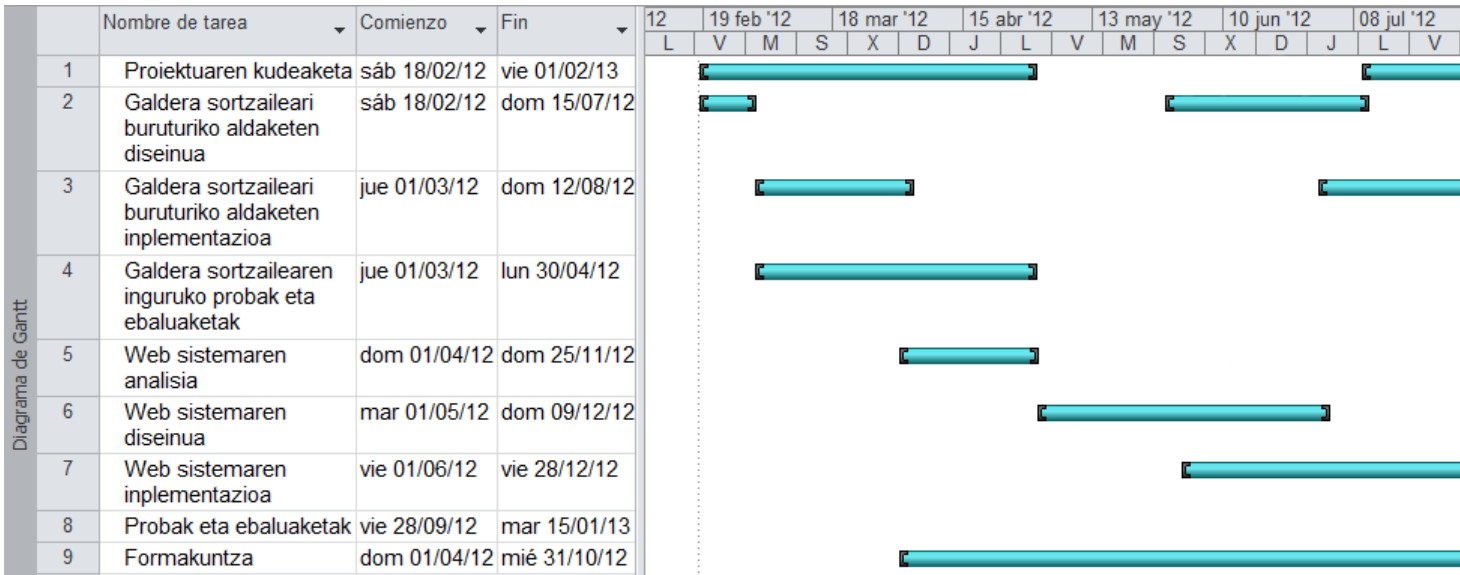
7.1	Birplanifikazioa	132
7.2	Atzerapenak eta denbora desbideraketak	135

Atal honetan proiektuaren helburu dokumentatuan egindako estimazioak eta errealitatean proiektuaren bizi zikloak bereganatu dituen epeak alderatuko dira.

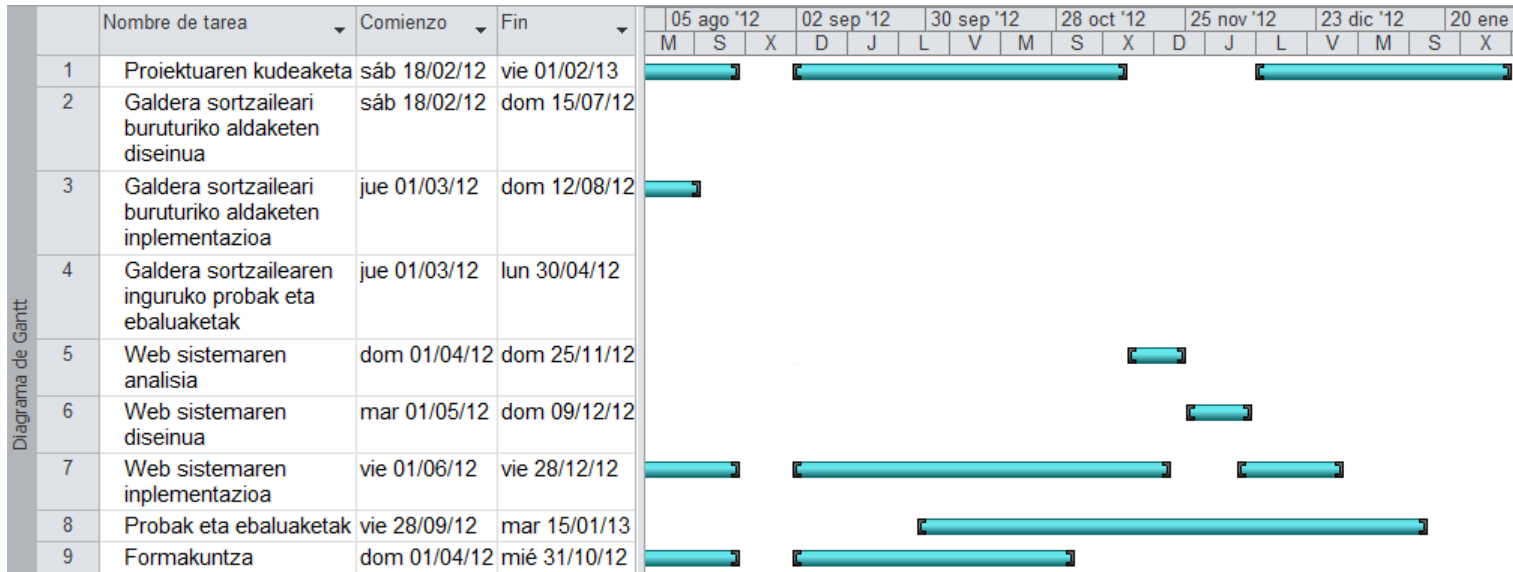
7.1 Birplanifikazioa

Proiektuaren bizi zikloaren hasieran ataza guztiak finkatu genituen eta hauek burutzeko behar izango genuen denbora estimatu genuen. Hala ere, hasiera batean finkatutako epemugak betetzen saiatu garen arren, arrazoi desberdinak direla eta, epe batzuk luzatu egin behar izan dira. Esaterako, proiektuaren amaieran tajuzko ebaluaketa bat egiteko hastirik ez genuela konturatu ginen, eta, ondorioz, proiektua birplanifikatzeko beharra agertu zen. Izan ere, ebaluaketa egiteko hainbat aldaketa bururatu zitzaizkigun, eta dudarik gabe, aldaketa hauek jotako denbora mugak gaindituko zituztela argi zegoen. Hala eta guztiz ere, hasierako Gantt diagrama oinarritzat hartuz -ikus 2.3, 2.4 eta 2.5 irudiak-, oro har epeak errespetatu dira.

Hurrengo irudietan Gantt diagrama berriak azaltzen dira, bizi zikloak bereganatu dituen denbora tarte errealak islatuz. Irudietan ikus daitekeen modura, epeak luzatzearen ondorio nagusia proiektuaren amaiera epea atzeratzea izan da. Izan ere, hasiera batean abendurako bukatuko genuela uste izan arren, azkenean, atzeratu behar izan da.



Irudia 7.1: Birplanifikazioaren Gantt diagrama.



Irudia 7.2: Birplanifikazioaren Gantt diagrama.

Gantt diagrama berriak kontuan hartuz, argi dago hasieran egindako estimazioekin guztiz bat ez datozela. Esaterako, estimazioetan proiektuaren kudeaketa bizi ziklo osoan zehar mantentzea aurreikusten genuen, aitzitik, lana burutzen joan bitartean uneoro hau mantentzea ezinezkoa izan da, lan karga astuna izan baitugu ataza batzuekin; ondorioz, atal asko dokumentatzean zirriborroak baino ez genituen eta kodera bueltatzea beharrezkoa izan da hainbat diagrama eta atal osatzeko. Honek guztiak, azken uneko dokumentazio lanak areagotu ditu.

7.2 Atzerapenak eta denbora desbideraketak

Azpiatal honetan proiektuaren hasieran egindako denbora estimazioa errealitatean behar izan den denbora kopuruarekin alderatuko dugu. Honetarako, proiektuaren helburu dokumentatua erabilitako taula berrerabiliko dugu. Hurrengo taulan orduen arteko aldea ikus daiteke, bigarren zutabean egin genituen estimazioak azaltzen dira, eta hirugarrenean, aldiz, errealitatean behar izan diren denbora tartekak.

Jotako orduak ikusita, estimatutako ordu kopurua errealitatean behar izan denetik gertu dagoela esan dezakegu. Oro har, desbideraketak atazaz ataza aztertzen baditugu ez dira hain nabariak izan. Arazo nagusienak proben eta ebaluaketen atalean aurki genitzake, bertan, estimatutako denborearen bikoitza baino gehiago behar izan baita. Desbideraketa honen atzean arazo desberdinak egonik ere, garrantzitsuenak ondokoak dira: lehenik eta behin, bukaeran diseinatutako ebaluazioa hasiera batean espero zena baino konplexuagoa izan da eta, gainera, ezusteko eta azken uneko hainbat aldaketa ekarri ditu, esaterako: hainbat funtzionalitate eta baliabide analizatu, diseinatu eta inplementatu behar izan dira, hala nola, galderak erantzuteko interfazean aurkitzen den laguntza baliabidea ia ia guztiz birdiseinatu behar izan da.

Atazak eta azpiatazak	Estimazioak	Errealitatea
Proiektuaren kudeaketa eta memoria	138 ordu	147 ordu
Bileren antolakuntza eta gauzatzea	40 ordu	30 ordu
Proiektuaren kontrola	5 ordu	4 ordu
Planifikazioa eta kudeaketa	8 ordu	8 ordu
Proiektuaren formalizaziorako memoria	75 ordu	95 ordu
Aurkezpen publikorako materiala eta prestaketa	10 ordu	10 ordu
Analisisa	17 ordu	17 ordu
Informazioaren bilketa eta aplikazioen analisisa	6 ordu	5 ordu
Aplikazioaren arkitektura	3 ordu	3 ordu
Helburuen zehaztapena	4 ordu	3 ordu
Arriskuen analisisa	1 ordu	1 ordu
Erabilpen kasuen diagramak	3 ordu	5 ordu
Diseinua	44 ordu	46 ordu
Galdera sortzaileari buruturiko aldaketen diseinua	18 ordu	14 ordu
Datu basearen diseinua	3 ordu	2 ordu
Web aplikazioaren diseinua	23 ordu	30 ordu
Inplementazioa	185 ordu	196 ordu
Software garatzeko teknologien aukeraketa	5 ordu	3 ordu
Softwarea garatzeko prozesua definitu	3 ordu	2 ordu
Galdera sortzaileari buruturiko aldaketen implementazioa	25 ordu	40 ordu
Datu basearen eraikuntza	2 ordu	1 ordu
Web aplikazioaren eraikuntza	150 ordu	150 ordu
Probak eta ebaluaketak	14 ordu	30 ordu
Ebaluaketara egokitzeko aldaketak	0 ordu	15 ordu
Formakuntza	35 ordu	20 ordu
Ataza eta azpiataza guztien batura	433 ordu	456 ordu

Taula 7.1: Proiektuaren denbora estimazioak vs errealitatea.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

8 Kapituluia

Ondorioak eta etorkizuneko lana

Gaien Aurkibidea

8.1 Ondorioak	138
8.1.1 Proiektuaren ondorioak	139
8.1.2 Ondorio pertsonalak	140
8.2 Etorkizuneko lana	141
8.2.1 Epe motzean egitekoak	141
8.2.2 Epe luzean egitekoak	142

Azken kapitulu honetan proiektuari dagozkion kontuak azalduko dira, hauen artean: ondorio orokorrak eta ondorio pertsonalak. Honetaz gain, implementatu gabeko eta proiektuaren etorkizunerako interesgarriak izan daitezkeen hainbat zeregin azalduko dira, gai guztiak orrazteko denborarik izan ez baita.

8.1 Ondorioak

Proiektu honetan euskaraz idatzitako testuak lantzeko erabilgarria den web-gune bat garatu da, erabiltzaileek igotako testuetatik abiatuta galderak automatikoki sortu eta material gisa eskaintzeko gai dena. Gainera, irakaskuntza/ikaskuntza arloan erabiltzen den metodologia berri baten bideari jarraitzen diona, kooperazioari eta elkar lanari alegia. Hau ahalbidetzeko, garatutako sistemak fitxategien eta ariketen partekatzea oinarritzat hartzen du, eta, honetaz gain, erabiltzaileen artean iritziak, balorazioak eta mezuak partekatzeko aukera ere ematen du.

Gainera, hezkuntza arloko aplikazio gehienek eskaintzen dituzten ikastaro itxien alternatiba gisa har daiteke aplikazioa, erabiltzaileek tutore lanak hartu eta euren materiala partekatuz, zuzenduz eta osatuz, beste erabiltzaileen ikasketa prozesuan laguntzeko gai direlako; hau guztiagatik, aplikazioaren osotasuna sistema erdiatomatikoko gisa ikus daiteke, erabiltzaileak izango baitira automatikoki sortutako ariketetan dauden akatsak zuzentzeko arduradunak. Honetaz gain, web aplikazioa automatikoki ariketaz hornitzen duen galdera sortzaile sistema ere aztertu da, lan poltsan garatutako lehen bertsioetik hasi eta hainbat hurbilpen linguistikoko berri erabiliz ahaltsuagoa den bigarren bertsio bateraino, honetarako, sarrera corpusaren tratamendua egokitu da, analizatzaile sintaktikoaren akatsak leuntzeko ahaleginak egin dira, hautagaien aukeraketa eta galdetzaileen esleipena egokitu da, eta markatze lengoaiak erabiliz galderak adierazteko formatuak diseinatu dira.

Halaber, sintagmen arteko urruntasuna neurtzeko bidean lehen urratsak eman eta garatutako proiektua ingurune errealetan ebaluatzeko saiakera desberdinak egin dira. Oro har, proiektuaren helburuak bete direla uste izan arren, burutu diren ebaluaketako emaitzek agindu modura, oraindik ere lan handia egiteko dago sistema funtzionala izaten hasteko. Dena den, ebaluaketatik abiatuta informazio oso baliagarria eskuratu da etorkizuneko lanak bideratzen hasteko.

Ondorioen kapitulua bi azpiataletan banatzea erabaki dugu: lehenik, proiektuarekin lotutako ondorioen atala eta, jarraian, ondorio pertsonalen atala. Ondorio pertsonalen azpiatalean proiektuaren bizi zikloan egindako lanak eragindako ondorioak azalduko dira. Proiektuaren ondorioetan, aldiz,

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

proiektuarekin zerikusia duten ondorio orokorrak azalduko dira.

8.1.1 Proiektuaren ondorioak

Proiektuaren helburu nagusia hizkuntza teknologiak bildu eta testuak lantzeko baliagarria den web aplikazioa garatzea izan da. Ez da software proiektu erraldoia izan, baina web teknologiaz baliatuta erabiltzailea pilare nagusi duen web 2.0 aplikazio bat garatu da. Era honetan, erabiltzaileen arteko elkarrekintza funtsezko da web aplikazioarentzat, ondorioz, erabiltzaile batzuk besteek egindako lanetan oinarrituko dira euskara lantzeko, beti ere, galderak automatikoki sortzen dituen sistemaren laguntzarekin.

Honetaz gain, aipatu beharra dago diziplinarteko proiektu bat izan dugula hau, zeinetan informatikariez gain hizkuntzalarien beharrak ere nabariak izan diren, bereziki galdera sorkuntzaren ildoan. Ondorioz, talde lanean eta espezialitate desberdinetako adituekin aritzeko aukerak eskaini dizkit proiektua egikaritu izanak, eskarmentu handia hortaz bada.

Halaber, proiektu ertain eta handiekin gertatzen den modura, kasu honetan ere, inplementazioari begira hartutako erabakiak -arkitektura eskemak, modulartasuna, aplikazioa geruzatan banatzea, MVC eredua eta metodologia arin bat jarraitzea- oso baliagarriak izan dira aplikazioa garatzeko, bete beharrei eta bat bateko aldaketei behar bezala erantzuteko malgutasuna eman baitigu. Gainera, aplikazioa modulu independentetan banatu denez, honek etorkizuneko lanak inplementatzeko erraztasunak emango dizkigu.

Hala ere, aurreko kasuan gertatu den modura, ondorio guztiak ez dira positiboak izan, izan ere, sarrera gisa erabiltzen ditugun analisisietan hainbat akats identifikatu ditugu -eranskinetako *Analisi linguistikoan arazoak* atalean (I.a) dokumentatu direnak- eta akats hauek gure sisteman eragin zuzena dutenez, lorturiko emaitzak okertu dituzte. Akats hauen guztien metaketaren ondorioz ebaluazioa eta erabilgarritasuna nabarmenki apaldu da, aplikazioaren ahalmena, alegia. Hala eta guztiz ere, euskararako aurki daitezkeen aplikazio kopuru mugatua dela eta, abantaila eta esker handikoa izan da tresna hauek erabilgarri eduki izana.

Kasu gehienetan gertatzen den modura, oraingoan ere denbora arazoak izan ditugu. Proiektuaren analisi, diseinu eta inplementazioak izan dituen denbora atzerapenak ez ezik, interfazearen ebaluaketak ere eragin handia izan du proiektua atzeratzeko garaian, izan ere, azken momentuko aldaketak ekarri ditu. Gainera, etorkizunerako alboratu ditugun ideiak direla, bai ebaluazio sakonegia prestatzeko intentzioa dela, baita txostena zein aurkezpena bukatzeko beharra dela; epemugak guregan etorri eta proiektua bukatzeko beharragatik, baliteke azken atal hauetan hankamotz geratu izana. Alabaina, aipatu beharra dago, hasieran aurreikusten ez ziren lan batzuk ere egin

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

direla, hala nola: erantzunen zuzenketa automatikoa itzulpengintzan erabiltzen diren metrikaz baliatuta, honen inguruan buruturiko lanak hastapenak baino izan ez badira ere.

8.1.2 Ondorio pertsonalak

Proiektuaren aurretik buruturiko lan poltsa erronka handia izan zen niretzat informatikako arlo guztiz ezezaguna bainuen hizkuntzaren prosesamendua, gainera, ikerkuntza munduan egiten nituen lehen urratsak izan ziren. Hone-taz gain, nabaria da NLP tresnak erabiltzeko garaiak euskararako baliabide kopuru murriztagoa dagoela -ingeleserako baliabideekin alderatuta batik bat- eta, honek, askotan geure eginbeharrak edo nahiak mugatu dituela, alabai-na, pixkanaka-pixkanaka sistema hobetzen genuen heinean eta arazoei aurre hartzen, emaitzak zer edo zer hobetzen ikusteak lanean jarraitzeko adina motibazio ematen zigun. Oro har, oso gustura egon nintzen lanean.

Lan poltsan garatutako sistemak hobekuntza ugari eskatzen zituen, eta, honetaz gain, sistema ebaluaketarik egin gabe geratu zen. Gainera, galde-rak sortzeko sistema testuak lantzeko erabilgarri izango zen ingurune batean integratu gabe geratu zen. Horregatik, hau guztia egiteko proiektua aukera ezin hobea izan da. Buruan genituen ideia guztiak bukatzeko hastirik egon ez bada ere, etorkizuneko lanetan dokumentatuta geratu dira, eta, oro har, hasiera batean finkatutako helburuak bete direla uste dut. Gainera, galderen sorkuntzan sakontzeko aukera izan dugu -egindako lana hastapenak baino ez diren arren-, arazoak eta etorkizunerako ideiak hobeto zehazten laguntzeko aukera paregabea izan dela uste dut.

Honetaz gain, proiektua Informatikan Ingenieritzako azken urtearekin ba-tera garatu dudanez, proiektua erronka handia bilakatu da, izan ere, aldi berean proiektuan zein irakasgaietan lan egiteak hasiera batean uste nuena baino konplexuagoa egin zait; eta momentuz behintzat, lan guztiak aurrera ateratzea lortu dut.

Dena den, ondorio guztiak ez dira positiboak izan, izan ere, galdera sor-tzaile sistemak webgunea gehiegi baldintzatzen duela uste dut. Jakina da testuak lantzea helburu duen web aplikazio batean ariketen kalitatea pre-miazkoa dela, baina, hala ere, nire ustetan eragin dion baldintzapena han-diegia izan da. Ondorioz, webgunea txukuna eta erakargarria suertatzeko intentzioarekin garatu bada ere, hainbat teknologia eta baliabide desberdin erabiliz garatu bada ere eta segurtasun neurriak zaintzeko diseinatu bada ere, hau guztia ezkutuan geratzen da, eta, azkenean, erabiltzaile arrunten iritzia sortutako galderen kalitatera bakarrik mugatzen da.

Azkeneko sentipenak, ordea, onak izan dira. Gogotsu aritu naiz lanean eta ibilbide honetan jarraitzeko gura dut, ikerkuntzan hain zuzen ere.

8.2 Etorkizuneko lana

Karrera bukaerako proiektua amaitutzat eman badugu ere, honen inguruan hainbat lan egin daitezkeela iruditzen zaigu, etorkizunean egin behar direnak, sistemaren errendimendua eta baliagarritasuna hobetzeko asmoz.

Egin beharreko lana bi multzotan banatzea erabaki dugu, interesgarri suertatuko delakoan: lehenengo multzoan epe motzean egingarri diren hobekuntzak bildu ditugu; eta bigarren multzoan aldiz, epe luzean egingarri direnak, astunago diren lanak alegia.

8.2.1 Epe motzean egitekoak

- Web aplikazioari lotutako hobekuntzak.

Erabiltzaileak kudeatzeko sistema konplexuago bat erabiltzea, era honetan, erabiltzaile bakoitzari pisu desberdinak esleituz, eta pisu honen arabera eragiketa gehiago edo gutxiago egiteko aukera eman - esaterako, pisu txikiagoa duten erabiltzaileen galderak zuzendu-

Galderak zailtasun mailaren arabera sailkatzea eta maila hauen arabera ere laguntza pertsonalizatua ematea.

Fitxategien azken eguneraketa datak kudeatzea.

Galderak erantzuteko atalari dagokionez eskolen metodologietara gehiago gerturatzen den interfazea eraikitzea. Era honetan, hasierako unetik laguntza eta emaitza zuzenak ikusteko aukera eman beharrean, gehigarri hauek akatsak egin ahala pantailaratzea. Esate baterako, akats bat egin ostean laguntza pantailaratuz eta bi akats egin ostean emaitza zuzena ikusteko aukera pantailaratuz, erabiltzailea ikasteko ahalegin handiagoa egitera bultzatuko genuke.

Web bilatzaileetan lortu dugun kokapena hobetu eta hitz gakoak hobeto finkatuz erabiltzaile gehiagorengana gerturatzea.

IXA taldearen boluntarioek proposatutako hobekuntzak inplementatzea.

- Proiektuarekin erlazionaturiko zereginak.

Proiektu honetan egindako lanak barneratuko dituen artikulua idaztea eta ondorengo bi *workshop*etako batean aurkezten saiatzea: *NAACL-HLT*¹ 2013 edo *NoDaLiDa workshop 2013*².

¹<http://www.cs.rochester.edu/tetreaul/naacl-bea8.html>

²<http://spraakbanken.gu.se/eng/Research/icall/2ndNLP4CALL>

8.2.2 Epe luzean egitekoak

- Galderen sorkuntzari lotutako hobekuntzak.
 - Anaforak ebaztea testua prozesatzean.
 - Esaldiak sinplifikatzea testua prozesatzean.
 - Semantika hiztegi edo sare konplexuagoak erabiltzea hautagaiak identifikatzeko fasean.
 - Patroietan oinarritutako galderak sortzea.
 - Patroi edo tripletetan -semantika ezaugarrietan- oinarrituta paragrafo mailako galderak sortzea.
 - Analizatzaile linguistikoan detektatu diren akatsak konpontzea.
 - Dependentzia sintaktikoak erabiltzea chunk mailan oinarrituriko QG sistema osatzeko.
- Web aplikazioari lotutako hobekuntzak.
 - Sortutako ariketak eta testuak audio formatuan ere eskura uztea, era honetan, irakurmena ez ezik, entzumena ere lantzeko aukera emanez.
 - Testuak eta ariketak lantzeko euskalki desberdinak kontuan izatea.
 - Esaldi mailako laguntza ematera mugatu beharrean paragrafo mailako laguntza ematea.
 - Ariketurri eta, oro har, ariketa sortzaile gehiago integratzea.
 - Galderak erantzuteaz gain, mota gehiagoko ariketak eskaintzea. Ariketa eredu zabalagoa sortzea alegia.
- Galderak zuzentzeko moduluari lotutako hobekuntzak.
 - Esaldien urruntasun kontzeptuan gehiago sakontzea eta metrika egokiagoak erabiltzea.

9 Kapituluak

Bibliografia

Gaien Aurkibidea

9.1	Liburuak	145
9.2	Artikuluak	145
9.3	Interneteko erreferentziak	147

Bibliografia

9.1 Liburuak

- [1] A. Astigarraga, K. Sarasola, K. Gojenola, A. Soroa, *TAPE Testu-analisirako perl erremintak*. UEU.
- [2] J.R. Zubizarreta, *Softwarearen ingeniari-tza*.
- [3] R. Nixon, *Learning PHP, MySQL and JavaScript*. O'Reilly Media.
- [4] E. Castro, *HTML con XHTML y CSS*. ANAYA MULTIMEDIA.
- [5] M. K. Glass, Y. L. Scouarnec, E. Naramore, G. Mailer, J. Stolz, J. Gerner, *Desarrollo Web con PHP, Apache y MySQL*. ANAYA MULTIMEDIA.
- [6] A. B. Downey, *Think Stats Probability and Statistics for Programmers*. O'Reilly Media.
- [7] E. T. Ray, J. McIntosh, *Perl and XML*. O'Reilly Media.
- [8] C. Darie, B. Brinzarea, F. Chereches, M. Bucica, *AJAX and PHP: Building Responsive Web Applications*. Packt Publishing Ltd.
- [9] S. M. Burke, *Perl and LWP*. O'Reilly Media.
- [10] C. Wong, *Web Client Programming with Perl Automating Tasks on the Web*. O'Reilly Media.

9.2 Artikuluak

- [11] E. Agirre (1), X. Artola (1), A. Diaz de Ilarraza (1), G. Rigau(1), A. Soroa (1), W. Bosma (2), *KAF: Kyoto Annotation Framework*. IXA group, University of the Basque Country (1), Computational Lexicology & Terminology Lab, VU Amsterdam (2).

- [12] I. Aldabe, M. Lopez de Lacalle, M. Maritxalar, E. Martinez, L. Uria, *ArikIturri: An Automatic Question Generator Based on Corpora and NLP Techniques*. UPV/EHU.
- [13] I. Aldabe, M. Maritxalar, A. Soraluze, *Question Generation Based on Numerical Entities in Basque*. *AAAI Symposium on Question Generation*. UPV/EHU.
- [14] I. Aldabe, *Automatic Exercise Generation Based on Corpora and Natural Language Processing Techniques*. UPV/EHU.
- [15] X. Yao, G. Bouma, Y. Zhang, *Semantics-based Question Generation and Implementation*. Johns Hopkins University, University of Groningen, Saarland University.
- [16] A. M. Olney, A. C. Graesser, *Question Generation from Concept Maps*. University of Memphis.
- [17] A. Díaz de Ilarraza, A. Mayor, K. Sarasola, *Semiautomatic Labelling of Semantic Features*. UPV/EHU, 19th International Conference on Computational Linguistics.
- [18] M. Liu, R. A. Calvo, *G-Asks: An Intelligent Automatic Question Generation System for Academic Writing Support*. University of Sydney.
- [19] M. Heilman, N. A. Smith, *Good Question! Statistical Ranking for Question Generation*. Language Technologies Institute.
- [20] I. Aldezabal, M. Aranzabe, A. Díaz de Ilarraza, A. Estarrona, L. Uria, *EusPropBank: Integrating Semantic Information in the Basque Dependency TreeBank*. UPV/EHU, Computational Linguistics and Intelligent Text Processing pp. 60-73.
- [21] I. Aldezabal, *Aditz-azpikategorizazioaren azterketa. 100 aditzen azterketa zehatza, Levin (1993) oinarri harturik eta metodo automatikoak baliatuz*. UPV/EHU.
- [22] Kl. Kipper, A. Korhonen, N. Ryant, M. Palmer, *Extending VerbNet with Novel Verb Classes*. Fifth International Conference on Language Resources and Evaluation (LREC 2006).

9.3 Interneteko erreferentziak

- [23] W3C-ko webgune ofiziala, <http://www.w3.org>.
- [24] Apache-ko webgune ofiziala, <http://www.apache.org>.
- [25] PHP-ko webgune ofiziala, <http://www.php.net>.
- [26] MySQL-ko webgune ofiziala, <http://www.mysql.com>.
- [27] Emacs-ko webgune ofiziala, <http://www.gnu.org/software/emacs/>.
- [28] Visual Paradigm-ko webgune ofiziala, <http://www.visual-paradigm.com/>.
- [29] Eclipse-ko webgune ofiziala, <http://www.eclipse.org/>.
- [30] Kile repositoria, <http://kile.sourceforge.net>.
- [31] IXA taldeko webgune ofiziala, <http://ixa.si.ehu.es/Ixa>.
- [32] IXA taldeko Eihera aplikazioaren gunea, <http://ixa.si.ehu.es/Ixa/Produktuak/1273220198>.
- [33] IXA taldeko Eulia aplikazioaren gunea, <http://ixa.si.ehu.es/Ixa/Produktuak/1274694158>.
- [34] IXA taldeko Ixati aplikazioaren gunea, <http://ixa.si.ehu.es/Ixa/Produktuak/1273220525>.
- [35] IXA taldeko Morfeus aplikazioaren gunea, <http://ixa.si.ehu.es/Ixa/Produktuak/1274695175>.
- [36] Wikipediako webgune ofiziala, <http://www.wikipedia.org>.
- [37] Euskaltzaindiako webgune ofiziala, <http://www.euskaltzaindia.net/>.

I Eranskina

Eranskinak

Gaien Aurkibidea

I.a	Analisi linguistikoan arazoak	150
I.b	SQL injection-en adibide erraz bat	153
I.c	Galdera sortzaile sistemaren ebaluaketako emaitzak	154
I.d	Eskuliburua	159

I.a Analisi linguistikoan arazoak

Adibide honetan chunk mailan agertzen diren -eta ohikoenak ere badiren- bi akats mota aurki genitzake:

1. Chunk mailan 'Orduan' eta 'eta' hitzak desagertu dira
2. 87 eta 94 lerroetan ikus daitekeen modura 'bi' eta 'kaletarrek' hitzen analisiak bikoiztuta agertzen dira

```

1 <?xml version="1.0" encoding="UTF-8" standalone="no"?>
2 <KAF xml:lang="eu">
3   <text>
4     <wf wid="w321" sent="41">Orduan</wf>
5     <wf wid="w322" sent="41">,</wf>
6     <wf wid="w323" sent="41">bi</wf>
7     <wf wid="w324" sent="41">kaletarrek</wf>
8     <wf wid="w325" sent="41">dentista</wf>
9     <wf wid="w326" sent="41">bati</wf>
10    <wf wid="w327" sent="41">deitu</wf>
11    <wf wid="w328" sent="41">eta</wf>
12    <wf wid="w329" sent="41">hagina</wf>
13    <wf wid="w330" sent="41">ateratzeko</wf>
14    <wf wid="w331" sent="41">agindu</wf>
15    <wf wid="w332" sent="41">zioten</wf>
16    <wf wid="w333" sent="41">.</wf>
17  </text>
18  <terms>
19    <!-- Orduan -->
20    <term tid="t230" type="close" lemma="orduan" pos="C.LOT-LOK"
21      >
22      <span>
23    <target id="w321" />
24    </span>
25    </term>
26    <!-- bi -->
27    <term tid="t231" type="close" lemma="bi" pos="D.DET-DZH">
28      <span>
29    <target id="w323" />
30    </span>
31    </term>
32    <!-- kaletarrek -->
33    <term tid="t232" type="open" lemma="kaletar" pos="N.IZE-ARR"
34      case="ERG">
35      <span>
36    <target id="w324" />
37    </span>
38    </term>

```

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

```

37     <!-- dentista -->
38     <term tid="t233" type="open" lemma="dentista" pos="N.IZE-ARR
      >
39         <span>
40 <target id="w325" />
41     </span>
42 </term>
43     <!-- bati -->
44     <term tid="t234" type="close" lemma="bat" pos="D.DET-DZH"
      case="DAT">
45         <span>
46 <target id="w326" />
47     </span>
48 </term>
49     <!-- deitu -->
50     <term tid="t235" type="open" lemma="deitu" pos="V.ADI-SIN">
51         <span>
52 <target id="w327" />
53     </span>
54 </term>
55     <!-- eta -->
56     <term tid="t236" type="close" lemma="eta" pos="C.LOT-JNT">
57         <span>
58 <target id="w328" />
59     </span>
60 </term>
61     <!-- hagina -->
62     <term tid="t237" type="open" lemma="hagin" pos="N.IZE-ARR"
      case="ABS">
63         <span>
64 <target id="w329" />
65     </span>
66 </term>
67     <!-- ateratzeko -->
68     <term tid="t238" type="open" lemma="atera" pos="V.ADI-SIN"
      case="ABS">
69         <span>
70 <target id="w330" />
71     </span>
72 </term>
73     <!-- agindu -->
74     <term tid="t239" type="open" lemma="agindu" pos="V.ADI-SIN">
75         <span>
76 <target id="w331" />
77     </span>
78 </term>
79     <!-- zioten -->
80     <term tid="t240" type="open" lemma="*edun" pos="V.ADL">
81         <span>

```

```

82 <target id="w332" />
83 </span>
84 </term>
85 </terms>
86 <chunks>
87 <!-- bi kaletarrek -->
88 <chunk cid="c144" head="t232" phrase="NP" case="ERG">
89 <span>
90 <target id="t231" />
91 <target id="t232" />
92 </span>
93 </chunk>
94 <!-- bi kaletarrek dentista bati -->
95 <chunk cid="c145" head="t233" phrase="NP" case="DAT">
96 <span>
97 <target id="t231" />
98 <target id="t232" />
99 <target id="t233" />
100 <target id="t234" />
101 </span>
102 </chunk>
103 <!-- deitu -->
104 <chunk cid="c146" head="t235" phrase="VP">
105 <span>
106 <target id="t235" />
107 </span>
108 </chunk>
109 <!-- hagina -->
110 <chunk cid="c147" head="t237" phrase="NP" case="ABS">
111 <span>
112 <target id="t237" />
113 </span>
114 </chunk>
115 <!-- ateratzeko -->
116 <chunk cid="c148" head="t238" phrase="VP" case="ABS">
117 <span>
118 <target id="t238" />
119 </span>
120 </chunk>
121 <!-- agindu zioten -->
122 <chunk cid="c149" head="t239" phrase="VP">
123 <span>
124 <target id="t239" />
125 <target id="t240" />
126 </span>
127 </chunk>
128 </chunks>
129 </KAF>

```

I.b SQL injection-en adibide erraz bat

Pentsa dezagun web orrialde batean erabiltzailea erregistratzeko eremu bat dagoela. Ohikoa den modura, erabiltzailea aplikazioan sartzeko bere izena eta pasahitza sartu behar ditu. Bere izena edo pasahitza jakin ezean ezin- go luke sisteman sartu, noski, baina zerbitzariak login sistema gisa azpiko kontsulta erabiltzen badu arazoak ager daitezke:

```
SQLstatement = "SELECT * FROM users WHERE name = ' " + $izena + "' and password = ' " + $pasahitza + "' ;"
```

Galdera egin ostean emaitzaren bat badago kautotzea onartu den seinale da. Ikus dezagun zer gertatzen den erabiltzaileak izen gisa "Iñigo" eta pasahitz gisa "1234" idazten badu:

```
$izena = "Iñigo" eta $pasahitza = "1234"
```

```
SQLstatement = "SELECT * FROM users WHERE name = 'Iñigo' and password = '1234' ;"
```

Kasu honetan ez litzateke arazorik egongo eta kontsulta ongi burutuko litzateke. Ondoren, ikus dezagun zer gertatzen den erabiltzaileak izen gisa "" or 1=1; -" eta pasahitz gisa "" idazten badu:

```
$izena = "" or 1=1; -" eta $pasahitza = ""
```

```
SQLstatement = "SELECT * FROM users WHERE name = ' ' or 1=1; - ' and password = ' ' ;"
```

Beraz, ikusten den modura, edozein erabiltzailek izango luke aukera kautotzeko, erabiltzaile konturik izan gabe. Izan ere, kontsulta lehen ';' karakterea dagoen tokian amaituko litzateke datu base sistemetan '- ' ikurrak iruzkin baten hasierari erreferentzia egiten baitio.

Txostenean aipatu den modura, kode txertaketa metodo hau arriskutsua da eta aplikazioak honen aurrean babestu behar dira. Honetarako PHP API-ak eskaintzen duen *mysql_real_escape_string* funtzioa erabiliko da.

Ikus dezagun zer gertatzen den erabiltzaileak izen gisa "" or 1=1; -" eta pasahitz gisa "" idazten badu, baina kasu honetan aipaturiko funtzioa erabiltzen bada:

```
$izena = mysql_real_escape_string("" or 1=1; -"); eta $pasahitza = mysql_real_escape_string("");
```

```
SQLstatement = "SELECT * FROM users WHERE name = '\ ' or 1=1; - ' and password = ' ' ;"
```

Kasu honetan erabiltzaileak ez luke etekinik lortuko zeren berak idatziriko ' karakterea baztertua izango litzateke, zehazki aurretik '\ ' karakterea erantsi zaiolako.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

I.c Galdera sortzaile sistemaren ebaluaketa-ko emaitzak

Ondoren, ebaluaketa egiteko erabilitako markak azaltzen dira

```
1
2 Gramatika ebaluatzeko:
3
4 Z → Guztiz ondo
5 Z, [azalpena] → Gramatikalki ondo, baina hiztunaren aldetik
   ez ohikoa
6 ?, [azalpena] → Gramatikalki arraroa
7 O → Guztiz oker
8
9
10 Galdetzaileak ebaluatzeko:
11
12 Z, [azalpena] → Guztiz ondo
13 O, [azalpena] → Guztiz oker
14
15
16 Egokitasuna ebaluatzeko
17
18 BAI → Guztiz egokia
19 ? → Zalantzazkoa
20 EZ → Ez da egokia
21 EZ, zaila ebaluatzeko → gramatikalki gaizki daude, ez dira
   ulertzen eta ezin da ebaluazio jator bat egin
```

Ondoren ebaluaketa burutu zuen hizkuntzalariak eginiko zuzenka batzuk ikus daitezke. Lehendabizi, gramatikari dagozkion emaitzak ageri dira eta, ondoren, galdetzaileen esleipenarena.

1	
2	Emaitza = Estatu Batuetatik
3	Sortutako galdera = Nondik etorri zan 1918ko gripea antza danez?
4	oharrak = Z, baina ez da Euskara Batua
5	egokitasuna: BAI
6	
7	Emaitza = bi egunetara
8	Sortutako galdera = Nora jaso nuen faktura handik?
9	oharrak = O
10	egokitasuna: EZ
11	
12	Emaitza = %20raino
13	Sortutako galdera = Noraino areago dezake aire egokitua axolagabekeriaz erabiltzeak kontsumoa?
14	oharrak = Z
15	egokitasuna: BAI
16	
17	Emaitza = beren produktuak
18	Sortutako galdera = Zer da milioika kontsumitzaileterengana zuzentzea eskaintzeko fabrikatzaileentzat ezinbestekoa?
19	oharrak = O, luzea
20	egokitasuna: ?, ebaluatzeko zaila
21	
22	Galdegia = gurasoen etxetik
23	Sortutako galdera = Nondik izango duzu azkenik hanka egiteko aukera?
24	oharrak = Z
25	egokitasuna: BAI
26	
27	Emaitza = gizonarekin
28	Sortutako galdera = Norekin ezkontzen denean emakumea berak aukeratutako?
29	oharrak = O, aditz nagusia falta
30	egokitasuna: BAI
31	
32	Emaitza = haltzak
33	Sortutako galdera = Zerk ez du bihotzik?
34	oharrak = Z
35	egokitasuna: BAI
36	
37	Emaitza = ganera
38	Sortutako galdera = Nora dago arazoa eskoletako bihotz-bihotzean ?
39	oharrak = O
40	egokitasuna: EZ

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

41
42 Emaidza = gurasoen etxetik
43 Sortutako galdera = Nondik izango duzu azkenik hanka egiteko
 aukera?
44 oharrak = Z
45 egokitasuna: BAI
46
47 Emaidza = zeruraino
48 Sortutako galdera = Noraino doa horrela telefono faktura?
49 oharrak = Z
50 egokitasuna: BAI
51
52 Emaidza = batekin
53 Sortutako galdera = Zerekin alderatu nituen biak guri ateratako?
54 oharrak = O, zerbait falta da ateratako ondoren
55 egokitasuna: EZ
56
57 Emaidza = publizitatearen inguruan
58 Sortutako galdera = Non eutsi gaitasun kritikoari?
59 oharrak = O, ahozkoa
60 egokitasuna: BAI
61
62 Emaidza = maiztasun handienarekin
63 Sortutako galdera = Zerekin dira erabiltzen direnak hauek?
64 oharrak = O, direnak edo hauek biak batera ez
65 egokitasuna: BAI
66
67 Emaidza = kezka horrekin
68 Sortutako galdera = Zerekin atera nintzen kalera?
69 oharrak = Z
70 egokitasuna: BAI

1	Jatorrizko esaldia = Reala azkenaldian bestean bizkarretik bizi da
2	Emaitza = bestean bizkarretik
3	Galdetzailea = Nondik
4	Galdetzaile egokia/oharrak = Z
5	
6	Jatorrizko esaldia = sopolarrak Freireri gertatu zaiona nabarmendu zuen
7	Emaitza = Freireri
8	Galdetzailea = Zeri
9	Galdetzaile egokia/oharrak = O, nori
10	
11	Jatorrizko esaldia = batzuetan bankuan kredituak eskatuta
12	Emaitza = bankuan
13	Galdetzailea = Non
14	Galdetzaile egokia/oharrak = Z
15	
16	Jatorrizko esaldia = datu hauek CNN telebista kateak eman ditu
17	Emaitza = CNN telebista kateak
18	Galdetzailea = Zerk
19	Galdetzaile egokia/oharrak = ?
20	
21	Jatorrizko esaldia = zortzi hilabetera etorri da bigarren pelotazoa
22	Emaitza = zortzi hilabetera
23	Galdetzailea = Nora
24	Galdetzaile egokia/oharrak = Z
25	
26	Jatorrizko esaldia = adin horretatik aurrera haurdun geratzeko aukera erdira jaisten da
27	Emaitza = adin horretatik
28	Galdetzailea = Nondik
29	Galdetzaile egokia/oharrak = ?, esaldi horretan: 'adin horretatik aurrera'
30	
31	Jatorrizko esaldia = nire kexua agertu nion
32	Emaitza = nire kexua
33	Galdetzailea = Zer
34	Galdetzaile egokia/oharrak = Z
35	
36	Jatorrizko esaldia = honek erronka berriak dakartza webgileentzat
37	Emaitza = webgileentzat
38	Galdetzailea = Zerentzat
39	Galdetzaile egokia/oharrak = O, norentzat
40	
41	Jatorrizko esaldia = epoa txirrindularitza profesionalean izurrite baten modura zabaldu zan

- 42 Emaidza = izurrite baten modura
43 Galdetzailea = Nora
44 Galdetzaile egokia/oharrak = O
45
46 Jatorrizko esaldia = etxeraino hurreratu zitzaion medikua
47 Emaidza = etxeraino
48 Galdetzailea = Noraino
49 Galdetzaile egokia/oharrak = Z
50
51 Jatorrizko esaldia = gure etxe azpian txoko berezi bat dago
52 Emaidza = gure etxe azpian
53 Galdetzailea = Non
54 Galdetzaile egokia/oharrak = Z
55
56 Jatorrizko esaldia = elikagaiak hamar taldetan banatzen dira
57 Emaidza = elikagaiak
58 Galdetzailea = Zerk
59 Galdetzaile egokia/oharrak = O, aditzak erg ez
60
61 Jatorrizko esaldia = politikarekin gauza bera gertatzen zait
niri
62 Emaidza = niri
63 Galdetzailea = Nori
64 Galdetzaile egokia/oharrak = Z
65
66 Jatorrizko esaldia = hortik aurrera abenduaren erdialderaino
eguraldia erabat aldakorra izango da
67 Emaidza = abenduaren erdialderaino
68 Galdetzailea = Noraino
69 Galdetzaile egokia/oharrak = Z
70
71 Jatorrizko esaldia = ez dira barazkiak iturriko uretatik pasa
behar
72 Emaidza = iturriko uretatik
73 Galdetzailea = Nondik
74 Galdetzaile egokia/oharrak = Z

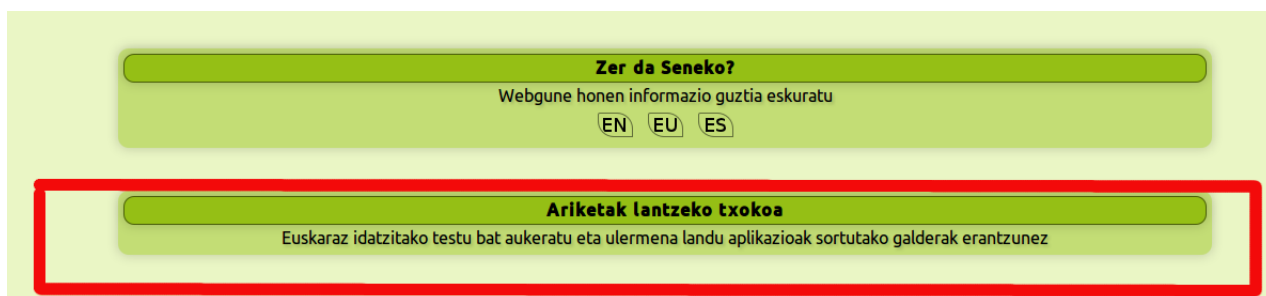
I.d Eskuliburua

Eskuliburu honetan web aplikazioa erabiltzen hasteko eta erabiltzen trebatzeko gidalerroak aurkitzen dira. Honetarako, bi jarduera proposatu eta pausoz-pauso azalduko dira:

1. Beste erabiltzaile baten fitxategi bat irakurri eta fitxategi horri dagozkion ariketak landu.
2. Sistemara igo dugun fitxategi bati galdera berriak gehitu.

Fitxategi bat irakurri eta ariketak landu

Lehenengo pausoa webgunearen sarrera orrialdera¹ nabigatzea izango da. Sarrera orrialdean hainbat funtzionalitate aurkezten badira ere -web aplikazioaren informazioa atzitu, erabiltzaile kontua sortu, pasahitza berreskuratu-, beste erabiltzaileen fitxategiak atzitzeko *Ariketak lantzeko txokoa* klik egin behar da.



Irudia I.1: Ariketak lantzeko txokoa klik egin.

Behin ariketak lantzeko txokoa izanik web aplikazioan atzigarri dauden fitxategi guztiak -fitxategi publiko guztiak- irakurri eta haiei buruzko informazioa eskuratu daiteke. Bilaketa errazteko asmoz eskuragarri dauden fitxategi guztien artean emaitzak filtratu ahal dira. Fitxategia aukeratu ondoren *irakurri* botoia sakatu.

¹<http://ixa2.si.ehu.es/seneko/>



Irudia I.2: Webgunearen sarrera orrialdea.

Izena	Gaia	Zailtasun maila	Balorazioa	Data	Egilea
Azeria eta saia.txt	Ipuina	L.H. (1.zikloa)	8 [iriziki ikusi]	2012/12/12	[irakurri]
Errege alferra.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe [iriziki ikusi]	2012/12/12	[irakurri]
Nire aitona Balentin.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/07	[irakurri]
gailurrak.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/09	[irakurri]
parrandatxo.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/09	[irakurri]
Manexen ipuina.txt	Ipuina	L.H. (3.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/10	[irakurri]
fitxategia Olatz ipuina lh2.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/10	[irakurri]
aitona baserritarra.txt	Antzerkia	L.H. (2.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/10	[irakurri]
NIREaitona.txt	Ipuina	L.H. (2.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/10	[irakurri]
sorginapirulak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
katuegalariak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
dantzariak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
lapurbihurriak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
lotiederrak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
baserritarrak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
ausartak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]
txakurabesleriak ipuina.txt	Ipuina	L.H. (1.zikloa)	baloratu gabe [iriziki ikusi]	2013/01/13	[irakurri]

Irudia I.3: Fitxategia aukeratu eta irakurri botoia sakatu.

Irakurri botoia sakatu ostean interfazean aukeratutako testua zabalduko da. Irakurketa amaitzean *Ariketak landu* botoia sakatu.

Nire_aitona_Balentin.txt - Erabiltzaileen balorazioa: baloratu gabe [itzuli] **Ariketak landu**

Nire aitona baserritarra da. Balentin deitzen da. Aurreko asteazkenean herrira jaitsi zen. Amonak dirua eman zion dentistarengana joateko eta hagina kentzeko. Aitona hagineko minnez zegoen.

Dentistaren kontsultara zihoala, gozotegi baten parean gelditu zen. Gozotegia pastelez eta gozokiz beterik zegoen: bonbak, kanutiluak, hojaldreak,... denetatik zegoen gozotegi hartan. Nire aitona Balentin oso gozozalea da eta pentsatzen hasi zen:

-Hagina ateratzen badut, dentistari ordaindu beharko diot eta ezin izango dut pastelik erosi. Dirua pasteletan gastatzen badut, berriz, hagineko minnez egongo naiz. Hain zuzen, hain komeria, hau!

Han zegoela, pastelei begira eta begira, erdi zorotuta, bi kaletar pasatu ziren handik eta burla eginez esan zioten:

- Zer dugu aitona? Pastel ederrak, e? Zenbat jango zenituzke?

Irudia I.4: Ariketak landu botoia sakatu.

Azkenik, fitxategiaren inguruko *arriketak lantzeko interfazea* pantailarazten da. Hiru eremu desberdinez osaturiko interfazea da: bata, galderen inguruko informazioa laburbiltzen duen taula; bestea, tresna erabilgarriak eskaintzen dituen taula; eta azkenik, galderak erantzuteko atala.

Seneko: galderak automatikoki sortuz testuak lantzeko aukera ematen duen aplikazioa

Ariketak lantzeko txokora itzuli

Ariketak lantzeko txokoa

Fitxategia : Nire aitona Balentia.txt

Hautaturiko fitxategiarl buruzko ariketen Informazioa Ezkutatu

Galdera kopurua guztira: 23 Landu nahi den galdera kopurua 23

Galdera motak :

Nor / Zer :	15	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norengandik / Nondik :	1	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norenganaino / Noraino :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norantz :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norengana / Nora :	2	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norekin / Zerekin :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Nori / Zeri :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norentzat / Zerentzat :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Nork / Zerk :	3	Hautatu lantzeko	<input checked="" type="checkbox"/>
Noren / Nongo :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Noren / Zeren :	0	Hautatu lantzeko	<input checked="" type="checkbox"/>
Norengan / Non :	2	Hautatu lantzeko	<input checked="" type="checkbox"/>

Filtroa ezarri Ezarri Denak hautatu Denak desmarkatu

Tresna lagungarriak

- Elhuyar itzultzailea
- Euskal Hiztegia
- Hiztegi Batua
- Hiztegi gehiago oraindik!
- Laguntza desaktibatu
- >Irakurri testua leiho berrian<--

1. Galdera Norengandik/Nondik zegoen bonbak kanutiluak hojaldreak gozotegi hartan?

Erantzuna

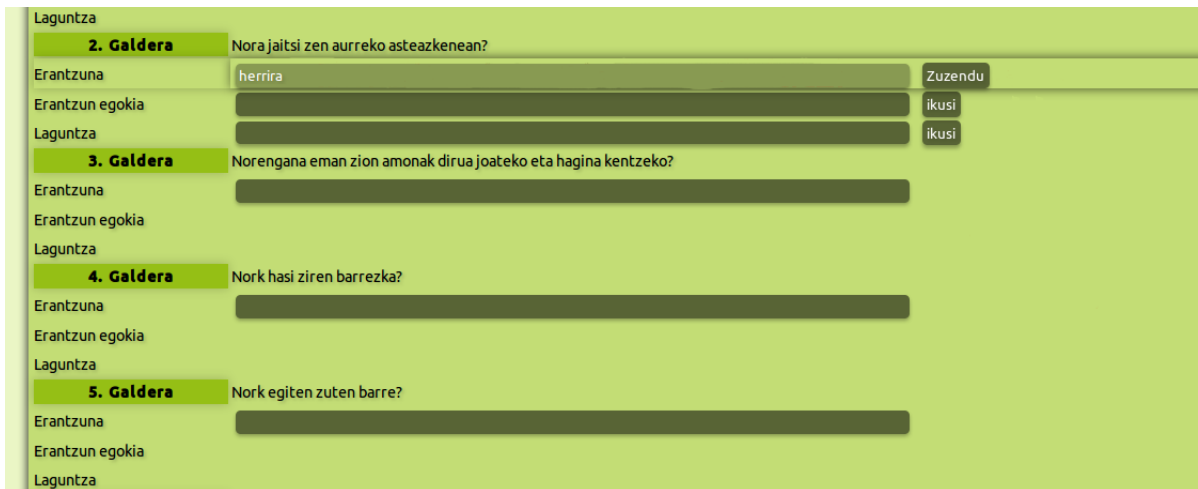
Erantzun egokia

Laguntza

2. Galdera Nora jaitsi zen aurreko asteazkenean?

Irudia I.5: Ariketak lantzeko interfazea.

Galderak erantzuten hasteko nahikoa da *azaltzen den laukitxoan klik egitea*, hau egitean lan ingurunea aktibatu egingo da eta tresna desberdinak agertuko dira. *Erantzunaren* atalean sistemak planteatutako galderaren emaitza idatzi behar da, honetarako, bai *laguntza* baita *erantzun egokia* baliabideak erabil daitezke. Baliabide hauek galderari buruzko pista bat eta erantzun egokia pantailaratzen dituzte hurrenez hurren, baliabide hauek erabiltzeko nahikoa da dagokion lerroko *ikusi* botoia sakatzea. Idatzitakoa zuzena den ala ez ikusteko, *zuzendu* botoia sakatu.

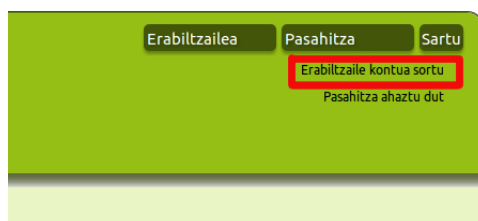


The image shows a screenshot of a questionnaire interface with a light green background. It contains five questions, each with a title, a question text, an answer input field, and a 'Laguntza' (Help) label. The first question is '2. Galdera' with the text 'Nora jaitsi zen aurreko asteazkenean?' and the answer 'herrira'. The second question is '3. Galdera' with the text 'Norengana eman zion amonak dirua joateko eta hagina kentzeko?'. The third question is '4. Galdera' with the text 'Nork hasi ziren barrezka?'. The fourth question is '5. Galdera' with the text 'Nork egiten zuten barre?'. Each question has a 'Laguntza' label and a 'Zuzendu' (Edit) button. The 'ikusi' (View) button is visible next to the answer fields for the first and second questions.

Irudia I.6: Galderak erantzuteko ingurunea.

Sistemara igo dugun fitxategi bati galdera berriak gehitu

Beste erabiltzaileen testuak lantzen ikasi ondoren, web sistema gure testuekin hornitzea da hurrengo urratsa. Era honetan, beste erabiltzaileek -igotako testuei partekatze publikoa emanik betiere- guk egindako lana oinarritzat hartuz egin ahalko dituzte ariketak. Ataza hau burutzeko bi baldintza hartu behar dira kontuan: alde batetik, erabiltzaile kautotuak bakarrik igo ditzakete fitxategiak web sisteman, bestetik, fitxategi bat editatzeko fitxategiaren jabea² izatea beharrezkoa da. Ondorioz, proposatutako ataza honekin has-teko erabiltzaile kontu bat sortu behar da. Honetarako, sarrera orrialdean atzigarri dagoen *Erabiltzaile kontua sortu* estekan klik egin.



Irudia I.7: Erabiltzaile kontua sortu.

Interfazeko pausoak jarraitzea nahikoa da erabiltzaile kontua sortzeko. Kontua guztiz gaitzeko *korreo elektronikoz jasotako mezuko estekan klik* egi-tea beharrezkoa da.

Pauso hauek burutu ostean web sisteman kautotu, honetarako, aurretik sortutako erabiltzailearen izena eta pasahitza idatzi eta sartu botoia sakatu. Hurrengo pausoa -proposatu den aurreko ariketaren modura- webgunearen sarrera orrialdera³ nabigatzea da.

²Fitxategiaren jabea fitxategia sisteman igo duen erabiltzailea da

³<http://ixa2.si.ehu.es/seneko/>

The screenshot shows a web interface for creating a new user. At the top left is the 'ixa Seneko' logo. The main content is a form titled 'Erabiltzaile berria sortu' (Create new user). The form contains the following fields and options:

- Izena (Name): text input field
- Abizenak (Surname): text input field
- Pasahitza (Password): text input field
- Errepikatu pasahitza (Repeat password): text input field
- Posta kodea (Postal code): text input field
- Korreo elektronikoa (Email): text input field
- Adina (Age): text input field
- Euskera maila (Basque level): dropdown menu with 'Altua' (High) selected
- Era **etikoan** jokatuko dut (I will play ethically): checkbox
- Onartu (Accept): button

At the bottom of the page, there is a footer: © Ixa taldea Webgunearen hasierako orrira itzuli

Irudia I.8: Erabiltzaile kontua sortu.

Sisteman kautotzean aurretik erabilgarri ez zegoen funtzionalitate berri bat agertuko da interfazean. Proposatutako ariketarekin jarraitzeko *Fitxategiak kudeatzeko txokoa*an klik egingo dugu. Txoko honetan fitxategi berriak igotzeko zein aurretik igotako fitxategien ariketen gainean edizioak egiteko baliabideak aurkitzen dira. Gainera, *Ariketak lantzeko txokoa*an gertatzen de modura *Fitxategiak kudeatzeko txokoa*an ere gure fitxategi pribatuen inguruko informazioa pantailaratzen da.



Irudia I.9: Funtzionalitate berri bat agertzen da kautotzean.

Interfaze honetan eskaintzen den fitxategiak igotzeko funtzionalitatea erabiliz, fitxategi bat igo, eta, honen ostean, interfaze honetara bueltatu. Fitxategi bat igotzeko nahikoa izango da fitxategiaren inguruan webguneak eskatzen dituen datuak betetzea eta *Igo* botoia sakatzea.



Inigo irten
Txatean parte hartu
 Profila aztertu
 Kontua ezabatu
 Erabiltzaileak kudeatu

Fitxategiak kudeatzeko txokoa

Nire fitxategien zerrenda - Ezkutatu

Izena	Gaia	Zailtasun maila	Balorazioa	Partekatzea	Data	Onartua			
proba.txt	Antzerkia	L.H. (1.zikloa)	baloratu gabe iriztiak ikusi	pribatua aldatu	2013/01/07	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
proba2.txt	Musika	Testu teknikoa	baloratu gabe iriztiak ikusi	pribatua aldatu	2013/01/08	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu
proba3.txt	Antzerkia	L.H. (1.zikloa)	baloratu gabe iriztiak ikusi	pribatua aldatu	2013/01/11	bai	Irakurri	Galderak zuzendu	<input type="checkbox"/> Ezabatu

Fitxategi bat aplikaziora igo

Izena

Gaia

Zailtasun maila

Partekatze modua

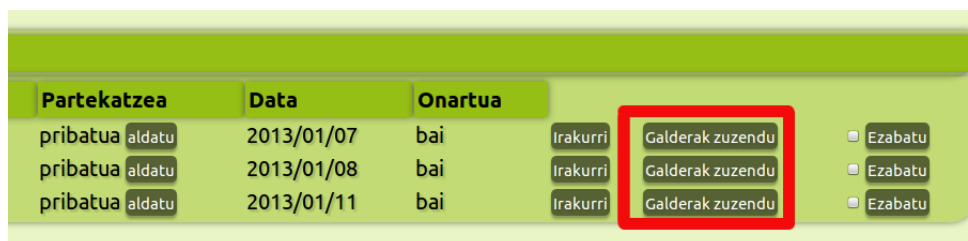
Fitxategia aukeratu (txt formatua, gehienez 25 kilo byte)

Era **etikoan** jokatuko dut

© Ixa taldea Webgunearen hasierako orrira itzuli

Irudia I.10: Fitxategiak kudeatzeko txokoa.

Ariketak editatzeko -bai automatikoki sortu direnak, baita aurreko edizio batean aldatutakoak- interfaze honetan bertan azaltzen den *Galderak zuzendu* botoia sakatu.



Irudia I.11: Galderak zuzendu botoia sakatu.

Galderak editatzeko agindua eman ondoren, fitxategiaren gaineko ariketak editatzeko interfazea pantailaratzen da, eta, bertan, ariketa bakoitzeko: galdera, erantzuna eta laguntza editatu daitezke. Honetaz gain, galdera bat ezabatu daiteke, honetarako, *Ezabatu galdera* botoia sakatuz. Galdera berriak sortzeko, aldiz, *Sortu galdera* botoia sakatzea beharrezkoa da. Edizioa amaitutakoan *Gorde* botoian sakatu behar da egindako aldaketak web sisteman gorde daitezen, aitzitik, egindako aldaketa guztiak galduko dira. *Ariketak lantzeko txokoa*n gertatzen den modura, *Fitxategiak kudeatzeko txokoa*n ere sortutako galderen inguruko informazio biltzen duen taula eta tresna lagungarriak biltzen dituen taula pantailaratzen dira.

Hautaturiko fitxategiari buruzko ariketen informazioa Ezkatatu

Galdera kopurua guztira: 2

Galdera motak :

Nor / Zer :	1
Noregandik / Nondik :	0
Norenganaino / Noraino :	0
Norantz :	0
Norengana / Nora :	0
Norekin / Zerekin :	0
Nori / Zeri :	0
Norentzat / Zerentzat :	0
Nork / Zerk :	0
Noren / Nongo :	0
Noren / Zeren :	0
Norengan / Non :	1

Tresna lagungarriak

Elhuyar itzultzailea
 Euskal Hiztegia
 Hiztegi Batua
 Hiztegi gehiago oraindik!
 -->Irakurri testua leiho berrian<--

Galdera	Zer erantzun zion azeriak?	Ezabatu galdera
Erantzuna	Hemen nago goitik behera datozen horiek zer ote diren begira	
Laguntza	Hemen nago goitik behera datozen horiek zer ote diren begira erantzun zion azeriak	
Kasua	Absolutiboa	
Galdera	Non zegoen azeria?	Ezabatu galdera
Erantzuna	Hemen	
Laguntza	Hemen nago erantzun zion azeriak	
Kasua	Inesiboa	
Galdera berria sortu	Galdera berri bat sortu nahi baduzu, aukeratu kasua eta sakatu botoia	Nor / Zer <input type="button" value="Sortu galdera"/>
Galderak gorde	Galderak zuzendu ondoren sisteman zuzenketak gorde nahi badituzu, sakatu ondorengo botoia (aurreko edukia galduko da)	<input type="button" value="gorde"/>
Berreskuratu	Akats bat egin baduzu, ez urduritu, ondorengo botoia sakatuz aurreko egoera berreskuratuko duzu	<input type="button" value="berreskuratu"/>

Irudia I.12: Galderak editatzeko interfazea.

