

▪ **Gradu Amaierako Proiektua** ▪

Software Ingeniaritza

Osasun-terminoak testuan ezagutzen

---

**Egilea: Ainhoa Elorz**

**Zuzendaria: Maite Oronoz Anchordoqui**

**Laguntzailea: Olatz Perez de Viñaspre**

2016- ekaina



# Laburpena

---

Txosten honetan aurkezten den lanaren helburua medikuntzarako euskarazko termino ezagutzaile bat egokitzea da. Tresna honek AnaMed\_eu du izena. Eginkizun nagusia testuan dauden osasun alorreko terminoak identifikatzea izango da, **SNOMED CT**-ren (*Systematized Nomenclature of Medicine – Clinical Terms*) eduki terminologikoa eta **Eustagger-lite** deituriko euskarazko analizatzailea erreferentziatzen izanik. Med\_en sistema abiapuntu nagusia izango da, tresna hori hobetuz eta euskarazko bertsioa sortuz. Azken horretarako, euskararen aberastasun morfologikoa kontutan izango da terminoak identifikatzeko garaian. Behin hori eginda, **Django** deritzon *frameworka* erabiliz sortutako interfaze batean idazten diren osasun-txostenetako terminoak kategorizatuak lortuko dira denbora errealean.

Hau da, memorian, euskarazko terminoen detektatzearen sorkuntza erakutsiko da Django *framework*-a erabiliz.

*Hitz gakoak: SNOMED-CT, medikuntza, Eustagger-lite, terminoak eta Django.*



# Gaien Aurkibidea

---

Sarrera eta motibazioa	1
Helburuak	4
Planifikazioa	6
<b>3.1 Proiektuaren irismena</b>	<b>6</b>
3.1.1 Betekizunen azterketa	7
3.1.2 Emangarriak	7
3.1.3 Mugak	7
3.1.4 Gantt diagrama	9
3.1.5 Produktuaren irismena	9
3.1.6 Lanaren Deskonposaketa eta Egitura (LDE)	10
<b>3.2 Denboraren planifikazioa</b>	<b>12</b>
3.2.1 Atazak	13
<b>3.3 Komunikazio-plana</b>	<b>14</b>
<b>3.4 Kalitatearen kudeaketa plana</b>	<b>14</b>
<b>3.5 Arriskuen kudeaketa-plana</b>	<b>17</b>
Teknologia	19
<b>4.1 Antzeko sistemak</b>	<b>19</b>
4.1.1 Text to SNOMED	20
4.1.2 FreelingMed	20
<b>4.2 SNOMED CT : osasun terminologia</b>	<b>21</b>
4.2.1 IHTSDO SNOMED CT Browser	23
<b>4.3 Analizatzailer linguistikoak</b>	<b>26</b>
4.3.1 CORENLP	27
4.3.1.1 Stanford CoreNLP interfazea Python-entzako	27
4.3.2 Eustagger-lite	28
4.3.2.1 NAF formatua	29
<b>4.4 Med_en sistema</b>	<b>31</b>
4.4.1 termMed_en	31
4.4.2 anaMed_en	32
<b>4.5 Python eta Django-a</b>	<b>32</b>
4.5.1 Programazio-lengoaia	32
4.5.2 Django	33
Analisia	40
<b>5.1 Arkitektura</b>	<b>40</b>
<b>5.2 Klase-diagrama</b>	<b>45</b>
<b>5.3 Osa tEUS egitura</b>	<b>47</b>
<b>5.4 Erabilpen-kasuak</b>	<b>49</b>
Inplementazioa	53
<b>6.1 Objektuen sorkuntza</b>	<b>54</b>

<b>6.2</b>	<b>Terminoak identifikatzen lewaren arabera</b>	<b>56</b>
<b>6.3</b>	<b>Datu-basea</b>	<b>59</b>
6.3.1	Baliabidearen aukeraketa	64
6.3.1.1	JavaScript Object Notation	65
<b>6.4</b>	<b>Wrapper-ra</b>	<b>71</b>
<b>6.5</b>	<b>Med_eu-en sorkuntza</b>	<b>73</b>
6.5.1	TermMed_eu:	73
6.5.2	AnaMed_eu:	74
<b>6.6</b>	<b>Django interfazea</b>	<b>77</b>
<b>6.7</b>	<b>Topatutako arazoak – konponbideak:</b>	<b>100</b>
6.7.1	Informazio lexikoa sortzen:	100
6.7.2	Terminoak txostenetatik detektatzen:	102
6.7.3	OsatEUS-en karaktere zehatzen ez deusaztaketa	105
	<b>Probak eta ebaluazioa</b>	<b>106</b>
<b>7.1</b>	<b>Probak</b>	<b>106</b>
<b>7.2</b>	<b>Ebaluazioa</b>	<b>116</b>
7.2.1	etiketazaileen arteko adostuna	120
	<b>Kudeaketa eta jarraipena</b>	<b>122</b>
<b>8.1</b>	<b>Burututako lana</b>	<b>122</b>
<b>8.2</b>	<b>Proiektua gauzatzeko plana</b>	<b>123</b>
8.2.1	Egutegia	124
8.2.2	Gantt diagrama	129
8.2.3	Lan-karga	130
<b>8.3</b>	<b>Komunikazioa</b>	<b>131</b>
<b>8.4</b>	<b>Kalitatea</b>	<b>131</b>
<b>8.5</b>	<b>Arriskuak</b>	<b>132</b>
<b>8.6</b>	<b>Mugak</b>	<b>133</b>
	<b>Ondorioak</b>	<b>136</b>
9.1	Ondorioak	136
9.2	Etorkizuneko lana	138
	<b>Bibliografia</b>	<b>139</b>
	<b>Erabilitako tresnak</b>	<b>144</b>
	<b>Bilerak</b>	<b>146</b>
	<b>Lehenengo Bilera</b>	<b>147</b>
	<b>Bigarren bilera</b>	<b>148</b>
	<b>Hirugarren bilera</b>	<b>149</b>
	<b>Laugarren bilera</b>	<b>150</b>
	<b>Bosgarren bilera</b>	<b>151</b>
	<b>Seigarren bilera</b>	<b>152</b>
	<b>Zazpigarren bilera</b>	<b>153</b>
	<b>Zortzigarren bilera</b>	<b>154</b>
	<b>Bederatzigarren bilera</b>	<b>155</b>
	<b>Hamargarren bilera</b>	<b>156</b>
	<b>Hamaikagarren bilera</b>	<b>157</b>
	<b>Hamabigarren bilera</b>	<b>158</b>
	<b>Hamahirugarren bilera</b>	<b>159</b>

Osasun txosten ereduak	<b>160</b>
<i>Sprint-ak</i>	<b>170</b>
Lehenengo sprinta	171
Bigarren sprinta	172
Hirugarren sprinta	172
Laugarren sprinta	173
Bosgarren sprinta	174
Seigarren sprinta	175
Zazpigarren sprinta	175
Zortzigarren sprinta	175
Bederatzigarren sprinta	176
Hamargarren sprinta	177
Hamabigarren sprinta	178
Hamahirugarren sprinta	178
Hamalauugarren sprinta	179
Eskuliburuak	<b>180</b>
Med_en-tzako eskuliburua	<b>180</b>
OsatEUS aplikazioaren eskuliburua	<b>184</b>

# Irudizerrenda

---

IRUDIA 3. 1: ESTIMAZIO-MUGARRI DIAGRAMA.....	8
IRUDIA 3. 2: ESTIMAZIO GANTT DIAGRAMA. ....	9
IRUDIA 3. 3: PROIEKTUAREN LAN DESKOPONSAKETA EGITURA (LDE) .....	11
IRUDIA 3. 4: USB PERTSONALEAN BALIABIDEAK GORDETZEKO FORMATUA. ....	18
IRUDIA 4. 1: SNOMED-CT-REN ONDASUN ZIKLIKOA.....	22
IRUDIA 4. 2: SNOMED-CT-REN 19 HIERARKIAK. ....	25
IRUDIA 4. 3: 3-HYDROXY-3-METHYLGUTARYL...-EN ARBASOAK. ....	25
IRUDIA 4. 4: ERLAZIOAK.....	26
IRUDIA 4. 5: EUSTAGGER-LITE-REN ARKITEKTURA. ....	29
IRUDIA 4. 6: HTML5-EN ELEMENTU SEMANTIKOAK.....	35
IRUDIA 5. 1: MED_EN SISTEMAREN HASIERAKO ARKITEKTURA. ....	42
IRUDIA 5. 2: MED_EN SISTEMAREN BUKAERAKO ARKITEKTURA.....	42
IRUDIA 5. 3: MED_EU SISTEMAREN BUKAERAKO ARKITEKTURA.....	44
IRUDIA 5. 4: BI SISTEMEN KLASE-DIAGRAMA.....	46
IRUDIA 5. 5: OSATEUS APLIKAZIOAREN EGITURA. ....	47
IRUDIA 5. 6: ERABILPEN-KASUAK.....	50
IRUDIA 6. 1: WRAPPER-AREN ITXURA.....	72
IRUDIA 6. 2: DIGESTIO-APARATUAREN INTERFAZEA HASIERAN.....	78
IRUDIA 6. 3: BARNE-MEDIKUNTZAREN INTERFAZEA BUKAERAN.....	79
IRUDIA 6. 4: ERABILTZAILERAK IDATZITAKO TESTUA LORTZEKO IRUDI GRAFIKOA.....	81
IRUDIA 6. 5: ANALIZATZAILERAK ETA WRAPPER_EUSTAGGER.PY-REN LINK BIRTUALAREN IRUDIKAPENA. ....	82
IRUDIA 6. 6: ANAMED_EU SISTEMAREN ERANTZUNA DJANGO-N LORTU.....	85
IRUDIA 6. 7: TERMINOAK EZ DIREN HITZEN ERAKUSKETA INTERFAZEAN.....	87
IRUDIA 6. 8: TERMINO ANITZEN DETEKTATZEA ETA ETIKETA ESLEIPENA.....	87
IRUDIA 6. 9:TERMINO ANITZAK ERA EGOKIAN LORTZEKO FUNTZIOAREN SASIKODEA.....	88
IRUDIA 6. 10: TERMINO ANBIGUOEN DETEKTATZEA.....	89
IRUDIA 6. 11: POPUP-AK HARTUKO DUEN ITXURA.....	91
IRUDIA 6. 12: POPUP LEIHOA.....	94
IRUDIA 6. 13: "HECES NEGRAS" TERMINOAREN ID-AREN BILAKETA.....	103
IRUDIA 7. 1: TERMINOEN ETIKETA MOTEN ESLEIPENA.....	114
IRUDIA 7. 2: "ASALDATUTA" TERMINOARI DAGOZKION SNOMED ID ETA FSN-AK.....	116
IRUDIA 8. 1: GANTT DIAGRAMA.....	129
IRUDIA 8. 2: MUGARRI DIAGRAMA.....	133
IRUDIA E. 1: STANFORD ZERBITZARIA MARTXAN.....	181
IRUDIA E. 2: LOGIN EGIN LEIHOA.....	184
IRUDIA E. 3: PASAHITZA BERREZARTZEKO LEIHOA.....	184
IRUDIA E. 4: ERABILTZAILERAK ERREGISTRATUAREN HASIERAKO LEIHOA.....	185
IRUDIA E. 5: BARNE MEDIKUNTZA ALTA TXOSTENA.....	185
IRUDIA E. 6: TERMINOEN DETEKTATZEA.....	186
IRUDIA E. 7: ASALDATUTA TERMINOAREN POPUP LEIHOA.....	186
IRUDIA E. 8: BARNE MEDIKUNTZA TXOSTENAREN LABURPENA.....	186



# Testu-kutxaren zerrenda

---

TESTU-KUTXA 3. 1: MUGARRIAK IDENTIFIKATZEKO EGITURA. ....	8
TESTU-KUTXA 3. 2: ESTIMAZIO-MUGARRIAK .....	8
TESTU-KUTXA 4. 1: "APOPTOSIS" KONTZEPTUAREN DESKRIBAPEN MOTAK.....	23
TESTU-KUTXA 4. 2: : "APOPTOSIS" KONTZETUAREN ERLAZIOA (BESTE KASU BATZUETAN BAT BAINO GEHIAGO EGON DAITEZKE). ....	26
TESTU-KUTXA 4. 3: "DOWN-EN SINDRO MEA" TERMINOA NAF FORMATUAN. ....	30
TESTU-KUTXA 4. 4: DJANGO-REN HASIERAKO FITXATEGIAK.....	33
TESTU-KUTXA 4. 5: CSSKO ID ETA CLASS IDENTIFIKATZAILEAK. ....	36
TESTU-KUTXA 6. 1: KONTZEPTU OBJETUAREN FUNTZIO HASIERATZAILEA.....	54
TESTU-KUTXA 6. 2: KONTZEPTU OBJETUAREN GET ETA SET FUNTZIOAK.....	54
TESTU-KUTXA 6. 3: HASIERATZE FUNTZIOARI DEITU.....	55
TESTU-KUTXA 6. 4: SET FUNTZIOAREN BIHURKETA.....	55
TESTU-KUTXA 6. 5: GET FUNTZIOAREN BIHURKETA.....	55
TESTU-KUTXA 6. 6: TERMINOEN LEMEI DAGOZKIEN SNOMED IDAK ESLEITU.....	57
TESTU-KUTXA 6. 7: LEMA BAKOITZAK EDUKIKO DUTEN SNOMED IDAK.....	57
TESTU-KUTXA 6. 8: TERMINO ANITZETATIK SNOMED ID LORTU.....	58
TESTU-KUTXA 6. 9: LEMEN ARABERA SNOMED ID-A LORTZEKO FUNTZIOA.....	58
TESTU-KUTXA 6. 10: INGELESEKO INFORMAZIO LEXIKOA DUEN FITXATEGIA SORTZEKO BEHAR IZAN DEN DENBORA.....	59
TESTU-KUTXA 6. 11.: EUSKARAZKO INFORMAZIO LEXIKOA DUEN FITXATEGIA SORTZEKO BEHAR IZAN DEN DENBORA.....	60
TESTU-KUTXA 6. 12: TSV FITXATEGIA SORTZERAKO AN DUEN ITXURA.....	61
TESTU-KUTXA 6. 13: TSV FITXATEGIA SORTU ONDORENGO ITXURA.....	61
TESTU-KUTXA 6. 14: PYTHONEN MYSQL ERABILTZEKO INPORTATU BEHAR DEN PAKETEA.....	62
TESTU-KUTXA 6. 15: MYSQL DATU BASEAREKIN KONEXIOA EGITEKO EMAN BEHAR DIREN URRATSAK. .....	62
TESTU-KUTXA 6. 16.: X OBJETUA F FITXATEGIAN GEHITU.....	62
TESTU-KUTXA 6. 17: F FITXATEGIA X OBJETUAN KARGATU.....	62
TESTU-KUTXA 6. 18: JSON-EN EGITURA.....	63
TESTU-KUTXA 6. 19: ANALYSIS_DIR ALDAGAI GLOBALAREN SORKETA.....	66
TESTU-KUTXA 6. 20: JSON FITXATEGIAREN KARGATEKA SELF.PRE_LOADED_ANALISYS_DICT ALDAGAIAN (BELTZEZ, AURRETIK AIPATUTAKO ALDAGAIK).....	66
TESTU-KUTXA 6. 21: 20.000 DESKRIBAPENERO FITXATEGIAN IDATZI.....	67
TESTU-KUTXA 6. 22: KONPROBATU EA D.TERM TERMINOA FITXATEGIAN DAGOEN.....	68
TESTU-KUTXA 6. 23:INFORMAZIO LEXIKOA DUEN FITXATEGIAREN ZATI BAT.....	68
TESTU-KUTXA 6. 24: IDAZKETAK EGITEKO BEHAR DEN DENBORA LORTU.....	69
TESTU-KUTXA 6. 25: ALERGIA TERMINO ANBIGUOA.....	71
TESTU-KUTXA 6. 26: ETIKETA BAKOITZARI BERE MULTZOA ESLEITU.....	72
TESTU-KUTXA 6. 27: : JSONFINALA ALDAGAI A BUELTATUKO DUENA (IKUSI F ERANSKINA).....	73
TESTU-KUTXA 6. 28: TS_EU.PY FITXATEGIA.....	74
TESTU-KUTXA 6. 29: EPONIMO LUZEENAREN LORKETA.....	74
TESTU-KUTXA 6. 30: EUSKARAZ EPONIMOAK EGOKI TRATATZEKO KODEA.....	75
TESTU-KUTXA 6. 31: SORTUTAKO WRAPPERA PAKETE BEZALA INPORTATU.....	75
TESTU-KUTXA 6. 32: JSON INFORMAZIOA LORTU.....	75
TESTU-KUTXA 6. 33: DJANGO FRAMEWORK-A BUKAERAN IZANGO DUEN ITXURA.....	77
TESTU-KUTXA 6. 34: ESPEZIALITATEA.HTML MOLDATU/GEHITUTAKO KODEA (BELTZEZ).....	80
TESTU-KUTXA 6. 35: ESPEZIALITATEA.HTML-EN EGINDAKO MOLDAKETAK CSS ARLOAN.....	81
TESTU-KUTXA 6. 36: HITZAKDETEKTATU BEHIN-BEHINEKO FUNTZIOA.....	82

TESTU-KUTXA 6. 37: JSONLORTU BEHIN-BEHINEKO FUNTZIOA .....	82
TESTU-KUTXA 6. 38: AJAX-EN ERABILERA TERMINOAK DETEKTATZEKO. ....	83
TESTU-KUTXA 6. 39: URLS.PY FITXATEGIA GEHITUTAKO LERROA. ....	83
TESTU-KUTXA 6. 40: AJAXETIK ANAMED_EU DEITZEKO PROZESUA. ....	84
TESTU-KUTXA 6. 41: VIEWS.PY FITXATEGIA. ....	84
TESTU-KUTXA 6. 42: ADENOKARTZINOMA-REN BI HIERARKIAK. ....	85
TESTU-KUTXA 6. 43: STRING - JSON KONBERTSIOA .....	86
TESTU-KUTXA 6. 44: 6.43 TESTU-KUTXAKO JSONLORTU (STRING) ETA JSO	86
TESTU-KUTXA 6. 45: IDENTIFIKATU ZEINTZUK EZ DIREN TERMINOAK. ....	87
TESTU-KUTXA 6. 46: "DOWN-EN SINDROME" TERMINO ANBIGUOAREN IDENTIFIKATZAILEAK.....	87
TESTU-KUTXA 6. 47: TERMINO ANITZEN ETA BAKUNEN KONTROLAKETA. ....	88
TESTU-KUTXA 6. 48: TERMINO ANBIGUOAK DEITURIKO FUNTZIOA. ....	89
TESTU-KUTXA 6. 49: "GAIXOTASUNAREN" SNO MED-ID-AK. ....	89
TESTU-KUTXA 6. 50: TERMINOEN ETIKETEN KOLORE ESLEIPENA. ....	90
TESTU-KUTXA 6. 51: POPUP HTML5 KODEA. ....	91
TESTU-KUTXA 6. 52: MODAL" CSS KLASEA. ....	92
TESTU-KUTXA 6. 53: "FADEIN" KEYFRAME-AK.....	92
TESTU-KUTXA 6. 54: "MODAL" CONTENT, HEADER, BODY ETA FOOTER CSS KODEAK. ....	93
TESTU-KUTXA 6. 55: SLIDEIN" KEYFRAMEAK. ....	93
TESTU-KUTXA 6. 56: X BOTOIAREN CSS KODEA.....	94
TESTU-KUTXA 6. 57: TERMINO A KLIK EGITERAKOAN POPUP-A AGERTZEKO. ....	95
TESTU-KUTXA 6. 58: POPUP-AREN ELEMENTUAK. ....	95
TESTU-KUTXA 6. 59: RADIOBUTTONEI ID ETA FSN AK LORTU.....	96
TESTU-KUTXA 6. 60: AJAX TEKNOLOGIA.....	97
TESTU-KUTXA 6. 61: URLS.PY FITXATEGIA. ....	97
TESTU-KUTXA 6. 62: SCT4TEXT() FUNTZIOA VIEWS.PY FITXATEGIAN. ....	97
TESTU-KUTXA 6. 63: SNOMED ID-AREN FSN-A.....	98
TESTU-KUTXA 6. 64: SCT4TEXT.PY TERMINALEAN ERABILITA.....	98
TESTU-KUTXA 6. 65: SCT4TEXT.PY-REN EDUKIA. ....	99
TESTU-KUTXA 6. 66: MOL DATUTAKO HITZAK. ....	101
TESTU-KUTXA 6. 67: EZABATUTAKO TERMINOAK. ....	101
TESTU-KUTXA 6. 68: RF1 FITXATEGIAN "EGINKARI BELTZ" LEMA GEHITU.....	103
TESTU-KUTXA 6. 69: KARAKTEREEN KONTROLA. ....	105
TESTU-KUTXA 8. 1: EGUTEGIAN ORDUAK IDAZTEKO FORMATUA.....	124
TESTU-KUTXA 8. 2: MUGARRIAK IDENTIFIKATZEKO EGITURA. ....	133
TESTU-KUTXA 8. 3: MUGARRIAK.....	134
TESTU-KUTXA E. 1: JAVA BERTSIOAREN ALDAK ETA .....	181
TESTU-KUTXA E. 2: KBP/ZERBITZARI DIREKTORIOKO EDUKIA.....	181
TESTU-KUTXA E. 3: STANFORD ZERBITZARIA MARTXAN JARTZEN .....	181
TESTU-KUTXA E. 4: : MEDTERM_EN ZERBITZARIA MARTXAN JARTZEKO KOMANDO A.....	181
TESTU-KUTXA E. 5: MEDTERM_EN MARTXAN.....	182
TESTU-KUTXA E. 6: TERMINOAREN ANALISIA EGITEKO KOMANDO A. ....	182
TESTU-KUTXA E. 7: "HEAD CACHE" TERMINOAREN ANALISIA.....	182
TESTU-KUTXA E. 8: MEDTERM_EN JASOTAKO ERANTZUNA .....	183
TESTU-KUTXA E. 9: STANFORD ZERBITARIA JASOTAKO ERANTZUNA .....	183
TESTU-KUTXA F. 1: PARSE FUNTZIOA.....	187
TESTU-KUTXA F. 2: TERMINOTEMP ETA ZUHAITZBIHURTU FUNTZIOAK.....	187
TESTU-KUTXA F. 3: NAF2JSON FUNTZIOA.....	187
TESTU-KUTXA F. 4: LABURTZAPENAK FUNTZIOA.....	187

## Grafikoen zerrenda

---

GRAFIKO 6. 1: DUMP() AGINDUAK ANALISIAK_EN.JSON FITXATEGIAN IDAZTEKO BEHAR DUEN DENBORA. ....	69
GRAFIKO 6. 2: DUMP() AGINDUAK ANALISIAK_EU.JSON FITXATEGIAN IDAZTEKO BEHAR DUEN DENBORA. ....	70
GRAFIKO 6. 3: AURKITUTAKO TERMINOEN KOPURUA .....	104
GRAFIKO 8. 1: HILABETEKO LAN ORDUAK .....	130

# Taulen zerrenda

TAULA 4. 1: HTML KODEAREN EREDU BAT.....	34
TAULA 4. 2: HTML ETA CSS-AREN EREDU BAT.....	35
TAULA 4. 3: HTML ETA CSS ID ETA CLASS ATRIBUTUEKIN EREDU BAT.....	37
TAULA 4. 4: JS ETA HTML-REN EREDU BAT.....	38
TAULA 4. 5: AJAX, JS ETA HTML-REN EREDU BAT.....	39
TAULA 6. 1: TERMINOA ETA LEMAREN ARTEKO EZBERDINTASUNA INGELESEZ.....	56
TAULA 6. 2: TERMINOA ETA LEMAREN ARTEKO EZBERDINTADUNA EUSKARAZ.....	56
TAULA 6. 3: BALIABIDEEN SARRERA TXIKIA.....	61
TAULA 6. 4: BALIBIDEEN AUKERAKETA.....	65
TAULA 6. 5: DUMP() AGINDUAK ANALISIAK_EN.JSON FITXATEGIAN IDAZTEKO BEHAR DUEN DENBORA.....	70
TAULA 6. 6: DUMP() AGINDUAK ANALISIAK_EU.JSON FITXATEGIAN IDAZTEKO BEHAR DUEN DENBORA.....	70
TAULA 6. 7: HIERARKIAK ETA HAUEN LEHENTASUNAK.....	85
TAULA 6. 8: AURKITUTAKO TERMINO KOPURUA.....	104
TAULA 6. 9: EUSTAGGER-LITEK ONARTZEN EZ DITUEN KARAKTEREAK.....	105
TAULA 7. 1: LENENENGO TXOSTENEAN LEHENENGO SAIAKERAN DETEKTATUTAKO TERMINOAK.....	110
TAULA 7. 2: BIGARREN TXOSTENEAN LEHENENGO SAIAKERAN DETEKTATUTAKO TERMINOAK.....	111
TAULA 7. 3: LENENENGO TXOSTENEAN BIGARREN SAIAKERAN DETEKTATUTAKO TERMINOAK.....	112
TAULA 7. 4: BIGARREN TXOSTENEAN BIGARREN SAIAKERAN DETEKTATUTAKO TERMINOAK.....	113
TAULA 7. 5: LENENENGO TXOSTENEAN HIRUGARREN SAIAKERAN DETEKTATUTAKO TERMINOAK.....	113
TAULA 7. 6: BIGARREN TXOSTENEAN HIRUGARREN SAIAKERAN DETEKTATUTAKO TERMINOAK.....	113
TAULA 7. 7: ETIKETA KOLOREA - HIERARKIA BALIOKIDETASUNA.....	115
TAULA 7. 8: KIRURGIA OROKORREKO ANAMED_EU GAIZKI DETEKTATUTAKO TERMINOAK 1. EBALUATZAILEAREN ARABERA.....	117
TAULA 7. 9: KIRURGIA OROKORREKO ANAMED_EU GAIZKI DETEKTATUTAKO TERMINOAK 2.EBALUATZAILEAREN ARABERA.....	119
TAULA 7. 10: ANESTESIOLOGIA ETA BERPIZTEKO ANAMED_EU GAIZKI DETEKTATUTAKO TERMINOAK 2.EBALUATZAILEAREN ARABERA.....	120
TAULA 7. 11: BI ADITUEN ETIKETATZAILEEN ARTEKO ADOSTASUNA.....	120
TAULA 8. 1: SORTUTAKO DESBIDERAKETAK.....	123
TAULA 8. 2: PROIEKTUAREN ATAZAK ETA ZENBAKIAK.....	124
TAULA 8. 3: ATAZA ETA ORDUA ESLEITZEKO ESKEMA.....	124
TAULA 8. 4: URTARRILEAN ETA OTSAILEAN EGINDAKO LANA.....	125
TAULA 8. 5: MATXOAN ETA APIRILEAN EGINDAKO LANA.....	126
TAULA 8. 6: MAIATZEAN EGINDAKO LANA.....	127
TAULA 8. 7: EKAINAN EGINDAKO LANA.....	128
TAULA 8. 8: MUGARRIEKIN SORTU DIREN DESBIDERAKETAK.....	134
TAULA D. 1: SPRINT-AK GAUZATZEKO SORTUTAKO ATAZAK.....	171
TAULA D. 2: LEHENENGO SPRINT-A.....	171
TAULA D. 3: BIGARREN SPRINT-A.....	172
TAULA D. 4: HIRUGARREN SPRINT-A.....	173
TAULA D. 5: LAUGARREN SPRINT-A.....	174
TAULA D. 6: BOSGARREN SPRINT-A.....	174
TAULA D. 7: SEIGARREN SPRINT-A.....	175
TAULA D. 8: ZAZPIGARREN SPRINT-A.....	175
TAULA D. 9: ZORTZIGARREN SPRINT-A.....	176

TAULA D. 10: BEDERATZIGARREN SPRINT-A .....	177
TAULA D. 11: HAMARGARREN SPRINT-A. ....	177
TAULA D. 12: HAMABIGARREN SPRINT-A. ....	178
TAULA D. 13: HAMAHIRUGARREN SPRINT-A.....	179
TAULA D. 14: HAMALAU GARREN SPRINT-A. ....	179

# ***1***

---

---

## **Sarrera eta motibazioa**

Osakidetzako (Euskal Autonomia Erkidegoko osasun zerbitzuak eskaintzen dituen erakundeko) langileek eta gaixoek bi hizkuntzetan tratua emateko eskubidea (euskara eta gaztelania EAEko hizkuntz ofizialak dira) dute. Ehuneko handi batek gaixoekiko tratua euskaraz eduki arren, idazten dituzten osasun-txostenak gaztelaniaz soilik dira. Hori hainbat arrazoiengatik gerta daiteke: terminoen ezagutza eza, haien ikasketetan barneratutako kontzeptuak gaztelaniaz barneratu izana, ...

Beste arrazoi nagusi bat izan daiteke Osakidetzak duen sistema zentralizatua, hau da, langile batek idazten duena beste batek irakur dezake eta alderantziz (hau ikus dezakegu medikura joaten garen bakoitzean, medikuek gaixoaren historial berdina dute ospitale ezberdinetan egon arren). Beharbada, osasun-txosten hura aurretik euskaraz idatzita egongo balitz, euskara ondo ez

dakien (edo ez dakien) mediku batek ez du ondo ulertuko eta gaixoaren osasuna arriskuan jarri dezake (terminoak gaizki ulertuz, adibidez).

Osakidetzan lan egingo duten mediku berriak, besteak beste, Euskal Herriko Unibertsitateko (EHU/UPV) batzuk, euskaraz graduatuak izango dira eta hizkuntza horretan matrikulatuta egon arren, haiek egiten dituzten praktikak gaztelaniaz izan behar dira, ikasgaien ehuneko txiki bat euskaraz ikasi izanik (berriro errepikatzen dugu gradua euskaraz egiteko matrikulatu zirela, hau da, beraien erabakia euskaraz ikastea izan zela). Zorte gehien izan duten ikasleek klase teorikoen %65a euskaraz eman dute Donostian. Gasteizen aldiz, euskaraz matrikulatuta egon arren, klase teoriko guztiak gaztelaniaz eman behar izan dituzte (Fdez. Serna, R., 2016). Gradua euskaraz ikastea zaila bada, nola lortuko dute lan egitea beraiek aukeratu zuten hizkuntzan?

Hori guztia dela eta, Osakidetzak bere 2013-2019 eperako definitu duen Osakidetzan Euskararen Erabilera Normalizatzeko 2. Planan zera zehazten du: “ahalegin handiak egin behar dira historia klinikoak ere ele bietan idatzita egoteko; eta horretarako, nahitaezkoa da luze gabe azterketa sakon bat egitea, informazioa sortu eta ustiatzeko prozesuari eragiten dioten faktoreak ondo ezagutze aldera”.

2014ko urtean IXA taldeko Olatz Perez de Viñasprek eta Maite Oronozek 45 osasun langileei galdeketa bat pasa zieten hurrengo galderekin: txostenak edozein hizkuntzatan idazten dituztenean, zein zailtasun topatzen dituzten ezagutzea eta txostenak euskaraz idazteko aukera izanez gero, tresnak zein behar ase beharko lituzkeen (ortografia-zuzentzailea, osasun-arloko terminologia euskaraz, txantiloiak etab.) zerrendatu nahi izan ziren eta hauek izan ziren ondorioak:

- Hizkuntzaren erabilerari dagokionez, 41 langilek gaixoekin komunikazioa euskaraz egiten dute baina 5ek soilik idazten dituzte txostenak euskaraz.
- Galdetegian osasun-txosten bat euskaraz idaztera gonbidatuak izan ziren inkestatuak eta honako zailtasunak aurkitu zituzten: euskarazko terminologiaren falta (45etik 32 pertsonak), eredu falta (24 pertsonak) eta ohitura falta (15 pertsonak).
- Euskaraz idazteko beharrezkoak iruditzen zaizkien baliabideen artean galdetuta, lehentasun handiena eman zieten euskarazko

medikuntzaren inguruko terminologia bateratu eta osatu bati (45etik 39 pertsonetik), medikuntzara egokitutako ortografia-zuzentzaile bati (30 pertsonetik) eta idazketarako eredu edo txantiloak izateari (25 pertsonetik).

Galdetutakoen %89 prest agertu zen osasun-txostenak euskaraz idazteko tresna sortuko balitz erabiltzeko, gainerakoek, berriz, bere zalantzak agertu zituzten; hau da, beste herrialde elebidunen antzera (Belgika edota Luxemburgo, adibidez), bi hizkuntza ofizial dauden tokietan, osasun-txostenak bi hizkuntzetan izatea nahi du Osakidetzak.

Gorago esan den moduan, lan hau IXA taldean kokatzen da, osasun arloko testuen hizkuntzaren azterketa automatikoaren alorrean hain zuzen ere. Beraz, proiektu hau ikerkuntza-talde baten barruan landu da.

Hurrengo atalean ikusiko dugu produktuak izango dituen helburu nagusiak.



# 2

---

---

## Helburuak

Euskara normalizazio-fasean dagoen hizkuntza izanik, hainbat alorretan bere erabilera mugatua da, eta osasungintza ere alor horien barruan kokatu daiteke. Osakidetzako osasun-zentroetan euskaldun langile anitz aurkitu badaitezke ere (kontsultak euskaraz emateko prest daudenak bezalakoak), alor horretan dauden gabeziak direla medio (software programak, adibidez), euskara erabiltzeko arazoak dituzte.

Horri aurre egiteko, euskarazko SNOMED CT lexikografia iturriak eta Eustagger-lite euskarazko analizatzaile/lematizatzaileak (3. atala) erremientek gabezia horren zati handi bat ase dezakete, medikuntzako terminologia zabala eta euskaraz dauden hitzen analisi sakona jasotzen baitute. Beraz, euskara bezalako hizkuntza gutxitu baterako, SNOMED CT antzeko sare garrantzitsuaren bertsio bat izateak motibatzen du lan hau ere.

Beraz, gradu amaierako proiektu honen helburu nagusia osasun alorrean euskara erabiltzeko bidean urrats bat ematea da, honetarako,

euskarazko osasun-terminologia, eta osasunera egokitutako analizatzaile bat sortzean datza idazteko sistema batean txertatuz. Era honetan medikuntzarako termino ezagutzaile bat egokituko dugu.

Gainera, esan beharra dago erremienta honen sorkuntza Osakidetza erakundeko mediku askoren eta IXA ikerkuntza-taldearen interesak asetuko zituela

Lehen helburua bete ahal izateko bidean, Med\_eu bertsioa sortzea izango da Med\_en (3. Atala) erremintatik abiatuz. Horretarako, Med\_en (ingeleseko termino ezagutzailea) erremintaren analisisa egingo dugu (4. atala) eta bertsio honi hobekuntzak egingo dizkiogu bilaketak eta bere martxan jartzea azkarragoa izateko. Behin hori eginda, euskarazko bertsioa egiten hasiko gara aurretik eginda zegoena baliatuz. Hau da, proiektu hau egiten den heinean bigarren mailako helburu bat beteko dugu, Med\_en sistema hobetzearena, alegia. Gainera, aplikazio bat egingo dugu erabiltzailearentzako euskarazko terminoen detektatzea errazago ikusteko eta terminoen hainbat informazioa lortzeko.

Med\_en eta Med\_eu sistemak Olatz Perez de Viñasprek sortuak dira. Proiektu hauetan Med\_en soluzioa hobetu egingo dugu lehenik, eta Med\_eu bertsioa ia osotasunean sortuko dugu Med\_en-en bertsioaren egokitzapenak eginez (zehaztapenak proiektuan zehar).

Beraz, aurretik aipatu dugun moduan, memoria honetan medikuntzarako euskarazko termino ezagutzaile bat egokitzea izango da helburu. Helburu hauek era egokian betetzeko, proiektua sortzeko ikasleak sortu duen planifikazioaz hitz egingo da 3.atalean.

# 3

---

---

## Planifikazioa

Proiektua modu arrakastatsuan amaitu ahal izateko planifikazio bat garatu da proiektuaren beharrei kasu eginez. Atal honetan arrisku nagusiak, kalitatea edo denborari buruz hitz egingo da.

### 3.1 Proiektuaren irismena

Azpiatal honetan proiektuaren irismenari buruz hitz egingo dugu. Gehiago zehaztearren, aipatuko diren atalak proiektuaren betekizunen azterketa, emangarriak, mugarriak, produktuaren irismena eta Lanaren Deskonposaketa Egitura (LDE) izango dira.

### 3.1.1 Betekizunen azterketa

Proiektu honetan Med\_en sistema hobetu eta Med\_eu garatu nahi da geroan osasun-txosten eleanitzak eduki ahal izateko. Bigarren sistema garatu ahal izateko lehenengo sistema erreferentziatza hartuko dugu hau hobetu ondoren. Med\_eu bertsioa sortzeko euskarazko terminoek dituzten berezitasunak kontuan hartuko ditugu (euskararen izaera eta sintaxia kontuan eduki beharko ditugu). Behin hau eginda, OsatEUS aplikazioa garatuko da euskal terminoak identifikatzeko eta kalifikatzeko.

### 3.1.2 Emangarriak

Proiektuan zehar bi emangarri mota sortuko dira:

– **Aplikazioarekin zerikusia dutenak:** produktua bukatzerakoan lau emangarri sortuko dira. Arazorik ez sortzeko, emangarri guztiak GAP direktoriaren barnean egongo dira:

- bi hizkuntzetarako sistemak (zerbitzariak eta analizatzaileak direktorioan).
- sistemak erabiltzeko eskuliburuak (E eranskina).
- informazio lexikoa eskaintzen duten bi fitxategiak (euskaraz eta ingelesez, .json formatuan) (baliabideak direktorioan).
- OsatEUS aplikazioa (osasunTxostenak direktorioan).

– **Proiektuarekin orokorrean zerikusia dutenak:** bi emangarri izango dira:

- Proiektuaren memoria.
- Proiektuaren defentsarako gardenkiak.

### 3.1.3 Mugak

Proiektuaren mugarri nagusienak diagrama batean irudikatu dira. Estimazio hauek hasieran egindakoak dira eta ez dituzte islatzen benetako datak.

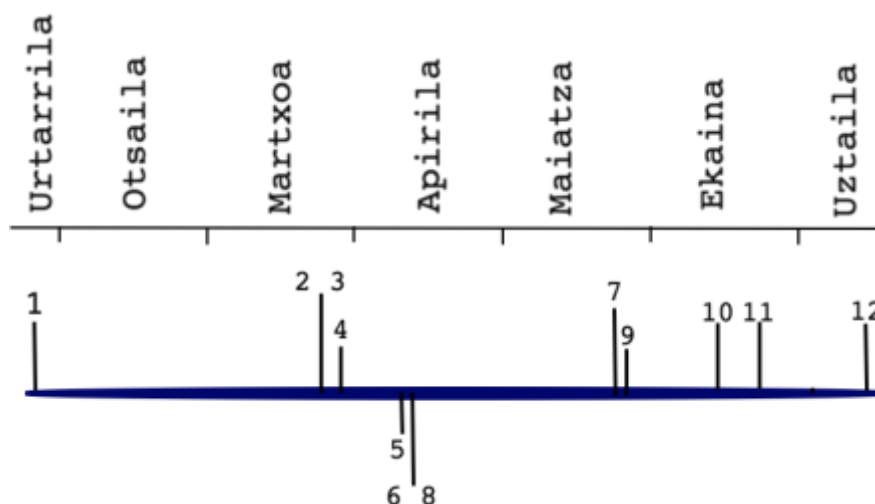
Diagrama garbiagoa izateko (ikusi 3.1 irudia), finkatuak izan diren mugarri bakoitzari zenbaki bat esleitu zaio. Zenbaki bakoitzari dagokion ataza

eta estimazio data ezarri zaio(ikusi 3.2 testu-kutxa). Estimazio-mugarri bakoitzak 3.1 testu-kutxako formatua izango dute.

Zenbakia.- Ataza: data estimatua

Testu-kutxa 3. 1: mugarriak identifikatzeko egitura.

Estimazio-mugarri diagrama:



Irudia 3. 1: Estimazio-mugarri diagrama

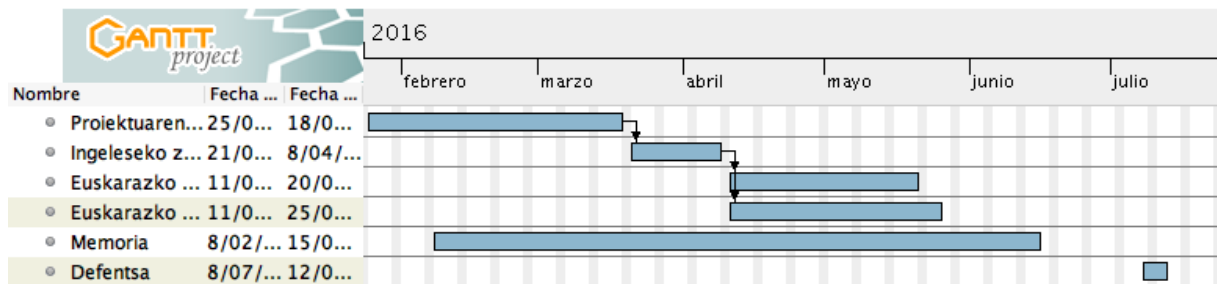
Estimazio-mugarri bakoitzari dagozkien atazak eta datak:

- 1.- Proiektuaren hasiera: urtarrilak 25
- 2.- Ingeleseko zerbitzaria hobetzen hasi: martxoak 20
- 3.- Memoriaren lehenengo zirriborroa: martxoak 20
- 4.- Planifikazioa: martxoak 28
- 5.- Ingeleseko zerbitzaria hobetzen bukatu: Apirilak 8
- 6.- Euskarazko zerbitzaria hasi: Apirilak 9
- 7.- Euskarazko zerbitzaria bukatu: Maiatzak 20
- 8.- Euskarazko bezeroa hasi: Apirilak 9
- 9.- Euskarazko bezeroa bukatu: Maiatzak 25
- 10.- Memoriaren azkeneko bertsioa: Ekainak 15
- 11.- Proiektuaren itxiera: Ekainak 23
- 12.- Defentsa: Uztailak 8-10-11

Testu-kutxa 3. 2: Estimazio-mugarriak

### 3.1.4 Gantt diagrama

Mugarriak kontuan hartuz, hurrengo Gantt diagrama<sup>1</sup> lortzea espero da (7.2 irudia):



Irudia 3. 2: estimazio Gantt diagrama.

Hainbat pakete handitan banatu da proiektua:

**Proiektuaren aurrekariak:** Hemen produktua egin ahal izateko hainbat terminoen barneraketa egingo dira (SNOMED-CT, AnaMed\_en, TermMed\_eu, ...). Gainera, Med\_en sistemaren klase-diagrama eta analisia egingo da.

**Ingeleseko zerbitzaria:** *StanFord* zerbitzaria hobetu egin da terminoen informazio lexikoa gordez fitxategi batean.

**Memoria:** Kurtsoan zehar egingo den proiektu guztiaren lana hemen islatuko da.

**Euskarazko Zerbitzaria (TermMed\_eu):** TermMed\_en-en euskarazko bertsioa sortuko da.

**Euskarazko bezeroa (AnaMed\_eu):** AnaMed\_en-en euskarazko bertsioa sortuko da.

**Defentsa:** Ikasleak proiektuaren zehar egindako lana aurkeztuko du.

### 3.1.5 Produktuaren irismena

<sup>1</sup> Proiektuen kudeaketan, proiektu edo lan-andana bateko bilakaera agertzen duen aurreikuspen bat da, lanen arteko bateragarritasunak eta ondorengotasunak kontuan harturik, azkenean proiektu osoaren iraupen minimoa emateko.

Produktu bakoitzak (ingeleseko eta euskarako sistemak) bezero eta zerbitzariz osatua egongo dira. Euskarako kasuan zerbitzaria eta bezero bakarra edukiko du eta ingelesekoan, aldiz, bi zerbitzari eta bezero bat.

Ingeleseko bertsioan aurretik eginda zegoena hobetuko da erantzun azkarragoa emanez. Horretarako, deskribapen guztien informazio lexikoa (*Stanford*-ek (4.3.1 azpitela) analisi sakonagoa egingo denez, denak kargatzen denbora asko behar du *TermMed\_en* zerbitzariak) datu-base batean gordeko ditugu (JSON fitxategi batean, alegia) *.txt* fitxategia ordezkatzuz. Gainera, terminoak lemaren arabera sailkatuko dira analisia egiterakoan (aurretik deskribapenaren arabera egiten ziren). Hau guztia euskararen bertsiorako oso lagungarri egingo zaigu. *Stanford* deritzon zerbitzaria *TermMed\_en* zerbitzariak pasatzen dizkien deskribapenen informazio lexikoa lortuko du. *TermMed\_en* zerbitzariak, aldiz, SNOMED CT-eko (4.2 azpitela) baliabideak kargatuko ditu bezeroak erabili ahal izateko.

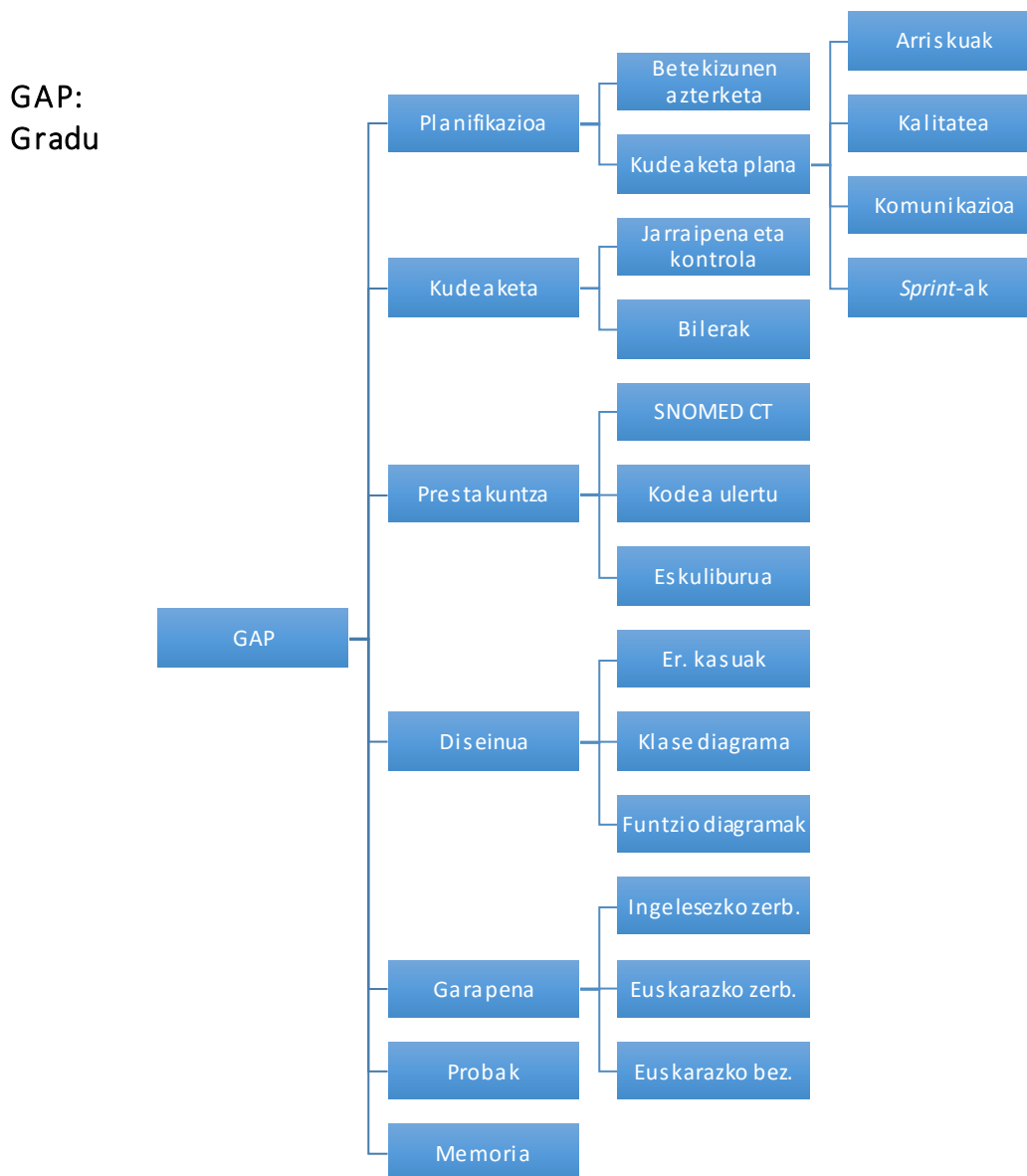
Euskarako bertsioa egiteko aldiz, *Eustagger-lite* (4.3.2 azpitela) deritzon analizatzaile linguistikoa erabiliko dugu deskribapenen, hau da, terminoen analisi lexikoa egiteko. Hau dena *wrapper* batean egingo da JSON formatuan terminoak bueltatzeko (*eustagger-lite*-k terminoen analisi lexikoa NAF formatuan bueltatzen baitu).

Behin hau eginda, gaizki tratatuak izan diren terminoak berriro analizatuko dira (adibidez, eponimoa den hitz mota lema bada, eponimoen tratamendua egin hauen atzikia ondo tratatzeko) euskararen morfologia ingelesekoa baino askoz konplexuagoa delako (atzizkiak kontuan hartu behar izango ditugu, besteak beste). Gainera, SNOMED CT erreferentzia izanik, terminoen analisiari "hierarkia", hitz bakoitzaren identitatea (*sctId*), etab. gehituko zaie. Hau egin ondoren, erabiltzaileek hainbat osasun-txosten idatzi ahal izango dituzte aurretik sortua izan den Django *framework* interfazearekin. Erabiltzaileak testu-kutxetan idazten dituen terminoak etiketatutako bueltatuko zaizkie (etiketa mota bakoitza kolore bat izango du). Terminoa (hitz bakuneko edo anitzekoa izanik) anbigua bada, lehenetsun gehiago duen hierarkiaren arabera koloreztatuko da. Erabiltzailea, etiketatuta izan den termino batean klik egiterakoan, *PopUp* leiho bat agertuko zaio, termino horren esanahi guztiak adieraziz *sctId* (SNOMED CT ID) eta *FSN* (*Full Specified Name*) elementuekin.

### 3.1.6 Lanaren Deskonposaketa eta Egitura (LDE)

Proiektu kudeaketaren ingurunean hierarkiatan deskonposatzen diren lan paketeak dira, hauen helburuak betetzeko asmoz. Lan pakete txikitzen banatzen da lana geroz eta errazago egiteko. Geroz eta lan pakete txikiagoak eduki errazagoak izango dira pakete hauek egiteko edo bukatzeko.

Proiektu honetan berebiziko garrantzia edukiko dute lan pakete hauek jarraipen eta kontrol egoki bat izateko (8. atala) eta atal honetan pakete bakoitzari espero diren lan orduak esleitu zaizkie.



### Amaierako Proiektua

- *Planifikazioa*: proiektua garatzeko plana.
  - o *Betekizunen azterketa*: proiektuak bete behar dituen ezaugarriak.

*Irudia 3. 3: Proiektuaren Lan Deskoposaketa Egitura (LDE)*



- *Kudeaketa plana*: kudeaketa egokia izateko plana. Hurrengo hauek kontuan hartuko dira.
  - Arriskuak.
  - Kalitatea.
  - Komunikazioa.
  - *Sprint*-ak.
- *Kudeaketa*: lanaren kudeaketa.
  - *Jarraipena eta kontrola*:
  - *Bilerak*: Zuzendari edo zuzendariordearekin edukitako bilerak.
- *Prestakuntza*: garapena ondo gauzatzeko gaitasunak lortzea.
  - *SNOMED CT*: analisia egin honen funtzionamendua ondo ulertzeko.
  - *Kodea ulertu*: aurretik eginda izan den lana barneratu.
  - *Eskuliburua*: aplikazioa martxan jartzeko eskuliburua sortu.
- *Diseinua*:
  - *Erabilpen-kasuak*.
  - *Klase-diagrama*.
  - *Funtzio-diagramak*.
- *Garapena*:
  - *Ingelesezko bi zerbitzariak*.
  - *Ingelesezko bezeroa*.
  - *Euskarazko zerbitzaria*.
  - *Euskarazko bezeroa*.
- *Probak*: sistemen funtzionamendu egokia hemen islatuko da hainbat proben bidez.
- *Memoria*: Proiektuari buruzko dokumentazioa.

## 3.2 Denboraren planifikazioa

Ataza bakoitzari eskainiko zaion denbora estimazioak dira. Estimazioak direnez, ez du zertan bukaerako emaitzarekin zer ikusirik izango.

### 3.2.1 Atazak

Proiektuari 300 ordu eskaintzea espero da. Hori kontutan hartuz, ondorengo estimazioak egin dira (LDE-aren paketeak erreferentziatzen hartu dira):

**(a) Planifikazioa: 19 ordu.**

- Betekizunen azterketa: 6 ordu.
- Kudeaketa-plana garatu: 13 ordu.
  - Arriskuen kudeaketa burutu: 2 ordu.
  - Kalitatea kudeaketa burutu: 5 ordu.
  - Komunikazio kudeaketa burutu: 2 ordu.
  - *Sprint*-ak: 4 ordu.

**(b) Kudeaketa: 20 ordu**

- Proiektuaren jarraipena eta kontrola: 8 ordu.
- Bilerak: 12 ordu.

**(c) Prestakuntza: 33 ordu**

- SNOMED CT (*Started Guide + Browser*): 17 ordu.
- Kodea ulertu: 15 ordu.
- Eskuliburua: 1 ordu.

**(d) Diseinua: 13 ordu**

- Erabilpen kasuak: 2 ordu.
- Klase diagrama: 5 ordu.
- Arkitektura: 6 ordu.

**(e) Garapena: 85 ordu**

- Sistemari ingelesez eusten dion zerbitzaria hobetu (*corenlp.py*) : 20 ordu.
- Sistemari euskaraz eusten dion zerbitzaria inplementatu (*ts\_eu.py*) eta *wrapperra* sortu (*wrapper\_eustagger.py*): 40 ordu.
- Sistemari euskaraz eusten dion analizatzailea inplementatu (*analizatzailea\_eu.py*): 25 ordu.

**(f) Proba funtzionalak: 20 ordu**

**(g) Memoria idatzi: 100 ordu**

GUZTIRA = a + b + c + d + e + f + g = 300 ordu

### 3.3 Komunikazio-plana

Proiektu honetan dauden interesatuak hiru dira:

<i>Ikaslea</i>	Ainhoa Elorz	<a href="mailto:aelorz003@ikasle.ehu.eus">aelorz003@ikasle.ehu.eus</a>
<i>Zuzendaria</i>	Maite Oronoz	<a href="mailto:maite.oronoz@ehu.eus">maite.oronoz@ehu.eus</a>
<i>Laguntzailea</i>	Olatz Pérez de Viñaspre	<a href="mailto:operezdevina001@ikasle.ehu.eus">operezdevina001@ikasle.ehu.eus</a>

Komunikazioaren zati handi bat posta elektronikoz burutuko da. Gainera, ikasle eta tutoreek noiz behinka bilerak egingo dituzte, eta hiruren agenden arabera, egun eta ordu bat finkatuko da. Bileretan ordura arte egindakoa aztertuko da, ikasleak aurkitu dituen arazoak azalduz, konponketak bilatuz eta ondorengo helburuak ezarriz. Ikaslea arduratuko da bileren laburpenak idazteaz, eta uneoro egongo dira laburpen horiek eskuragarri proiektuan eranskin moduan (B Eranskina). Olatz Perez de Viñasprek proiektuaren laguntzaile edo zuzendari ordearen papera hartuko du.

Kode guztia *GitHub*-eko biltegietan egongo da eta ikaslea kodea biltegira aldatetak egin ondoren igotzeaz arduratuko da. Horrela tutoreek kodea eskura edukiko dute edozein momentuan.

### 3.4 Kalitatearen kudeaketa plana

Proiektua kalitate maila bat ezarri diogu eta kalitate hura mantentzeko hurrengo hauek kontuan hartu dira:

#### 1.- Proiektuaren kudeaketa

Gutxienez hiru astetik behin *sprint*-ak (D Eranskina) finkatuko ditu ikasleak (zuzendariekin nahiz bera bakarrik) helburu batzuk finkatuz. Honen helburu nagusia lan-pakete batzuk finkatzea izango da proiektua aurrera eramateko. *Sprint*-ak proiektu guztian eragina izango du, hau da, ez produktuan bakarrik baizik eta proiektu guztian (memoria edo produktua, esate baterako).

Klaseak amaitu aurretik (aste santu aurretik) gutxienez hiru asteetan behin zuzendari edo laguntzailearekin komunikazioa burutuko da, emailez, *Skype* edo bilera batekin behar izanez gero. Horrela, ikasleak produktuaren egoeraren berri emango du (zailtasunak, moldaketak, egindako lan berria, ...) bide egokian doan ikusteko.

Klaseak amaitu ondoren eta proiektua entregatzeko epea hurbiltzen ari den heinean astean behin bilerak egingo dira.

## 2.- Garatutako produktuak

Produktuaren kalitateari dagokionez hiru zerbitzariak (euskarazko `MedTerm_eu` eta ingelesezko `MedTerm_en` eta `Stanford` direnak) bereziko arreta jarriko zaie hurrengo arrazoiengatik:

- Ingeleseko zerbitzariak (`corenlp.py` eta `ts_en.py`): oinarrizko zerbitzariak direlako eta hauek oinarri hartuta euskarazko zerbitzaria garatuko delako.
- Euskarazko zerbitzaria (`ts_eu.py`): produktua sortzeko zati garrantzitsu bat izango direlako. Beste hitzetan, proiektuaren bukaerako emaitza euskarako terminoak analizatzea izango da. Terminoen analizatzailea zerbitzari duen eta `wrapper_eustagger.py` `wrapperraren` beharra izango du.
- Ingeleseko bezeroa (`analizatzailea_en.py`): erabiltzaileak ingeleseko terminoaren analisia jasoko du.
- Euskarazko bezeroa (`analizatzailea_eu.py`): euskarazko terminoaren analisi sakon bat jasoko du euskarak dituen ezaugarriak kontuan izanik.
- `OsatEUS` aplikazioa: erabiltzaileek erabiliko duten aplikazioa delako terminoek etiketatu ahal izateko.

Gainera, zerbitzariaren kalitateak berebiziko garrantzia izango du terminoen analisiaren kalitatea hauen araberakoa izango baita. Bi bezeroek (`analizatzaile_eu` eta `analizatzaile_en`) kalitate eskaseko zerbitzariak erabilia analisi eskasak edukiko lituzkete eta ez lukete analisi egokia bueltatuko. Honengatik kalitate handiko zerbitzariak beharko ditugu.

Eskuliburu emangarria (E eranskina) esan den bezala Java bertsioa exekutatu beharko dugu (1.8 bertsioa) ingeleseko sisteman (euskarazko

sisteman ez dugu Javaren bertsio beharrik ez dugulako *Stanford* erabiltzen). Proba eta espero diren emaitzak memorian bertan idatziko dira (7. atala). Hainbat proba mota desberdin egingo dira, adibidez:

- Hitz bakarreko terminoekin.
- Hitz anitzeko terminoekin.
- Eponimoak dituzten terminoekin (Down, Park, Parkinson, ...)
- Karaktere espezialak erabiltzen dituzten terminoekin ('+', ':', gaztelaniak dituen karaktere bereziak, ...).
- Interfazean sartutako testuak
- ...

Hasiera batean probek huts egingo dute zerbitzari eta bezeroaren inplementazio eskasarengatik eta beharrezkoa den kodea gehituko zaio proba modu zuzenean atera. Behin probak huts egiten ez duenean kode zati hau gordeko da gerora akatsik gertatzen bada hau berregituratu ahal izateko. Ziklo hau behin eta berriz errepikatuko da proba guztiak ondo egin arte. Hau da, probak ez dira automatizatuak egongo eta erabiltzaileak egin beharko ditu.

### 3.- Memoria

Zuzendari eta zuzendari ordeari bere iritzia eskatuko zaie memoriaren sortze prozesuan.

Horrez gain, proiektuarekin loturarik ez duen pertsona baten iritzia ere jasoko da (informatika arloan jakintza duen pertsona bat izango da), batez ere ideia orokorrak ongi azaltzen diren jakiteko (teknologia, proiektuaren helburua, garapena, ...). Honen helburua honakoa da: proiektuan murgilduta dagoen batek testuingurua ongi ezagutuko du eta informazio guztia ez jaso arren, esanahi dena ulertuko du. Beharbada, puntu edo termino batzuk ez dira ongi azalduak izan eta horregatik loturarik ez duen pertsona baten iritzia nahi da.

Memoriaren tamaina egokitzat hartu da hauek 100-120 orrialde inguru izatea. Hala ere, hau ez da garrantzitsua izango kalitatea premiazkoa izango baita, hau da, ez da edonolako edukirik sartuko orri gehiago edukitzeko edo beharrezko informazioa kenduko finkatutako orrietara iristeko.

Akats ortografikoek eta erabilitako euskarak bereziki arreta handia izango du memorian.

## 3.5 Arriskuen kudeaketa-plana

Jarraian, proiektua kolokan jarri dezaketen arriskuak aipatuko dira. Zer eragin eduki dezaketen eta nola ekidin daitezkeen ere azalduko da. Eraginaren arabera (oso handia, handia, ertaina edo txikia) eta gertatzeko probabilitatearen arabera ordenatu dira:

### 1.- Produktua ezin amaitzea

**AZALPENA:** proiektua ez bukatzea interesatuek finkatu duten epean, ikasleak urtean eduki duen lan zamaren ondorioz.

**ERAGINA:** oso handia.

**GERTATZEKO PROBABILITATEA:** ertaina.

**KONPONBIDEA:** beranduago aurkeztea proiektua, ikasleak egunean lan ordu gehiago sartzea edo helburu gehigarriak ez betetzea eta hauek proiektuaren hobekuntza moduan aurkeztea.

### 2.- Denbora-estimazio okerrak egitea

**AZALPENA:** ataza bakoitza burutzeko denbora gaizki estimatzea.

**ERAGINA:** ertaina.

**GERTATZEKO PROBABILITATEA:** handia.

**KONPONBIDEA:** ataza bakoitzean pasatako denbora kontrolatuko da eta estimazioetatik urruntzen badoaz zergatia aztertuko da (8. atala). Estimazioak ez dira egokituko, hau da, hasierako estimazioak jarraituko dute ikusgarri eta benetan sartu diren ordu kopurua azalduko da hainbat grafiko bidez (asteroko ordu kopuruak, ataza bakoitzean espero ziren eta benetan sartutako orduak, asteroko lan karga, ...)

### 3.- Proiektuko baliabide guztiak galtzea

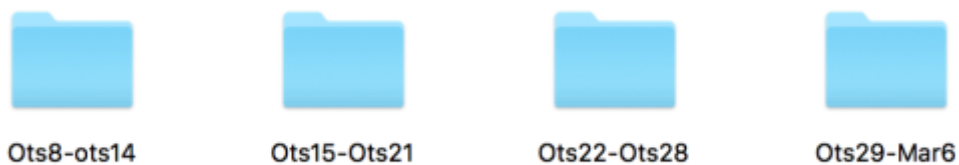
**AZALPENA:** hainbat baliabidek huts egitea eta proiektu guztia galtzea (ordenagailua apurtzea, sarea erortzea, ...).

**ERAGINA:** oso handia.

**GERTATZEKO PROBABILITATEA:** txikia.

**KONPONBIDEA:** dokumentazio eta kode guztia ikaslearen ordenagailu pertsonalean egongo da eta aurretik esan dugun bezala, kodearen kopia *GitHub*-en egongo da ere. Kodea, gainera, IXA taldeko zerbitzarietan aurkituko da ikaslearen kontu pertsonalean.

Ikaslearen USB pertsonalean dokumentazioaren eta kodearen kopiak egingo dira astero. Honela, aurreko bi baliabideek (GitHub eta ordenagailu eramangarria) porrot egiten badute beste kopia bat izango dugu. Hau da, astero sortu izan diren azken bertsioak gordeko dira karpetetan asteroko data jarritz, baldin eta aste horretan zeozer egin bada GAP-ean.



*Irudia 3. 4: USB pertsonalean baliabideak gordetzeko formatua.*

4. atalean, hainbat erraminta erabiliko ditugu erreferentziatzen produktua gauzatzeko, lehendik beraien funtzionamendua ulertuz.

# 4

---

---

## Teknologia

Atal honetan produktua egin ahal izateko zein teknologiaz baliatu garen eta zeintzuk baztertu ditugun azalduko dugu arrazoi justifikatuak emanez. Lehenik eta behin, aurretik sortua zeuden antzeko sistemei buruz hitz egingo dugu zer diren esanez eta zertarako balio duten azalduz.

### 4.1 Antzeko sistemak

Azpi atal honetan jadanik existitzen diren tresnak azalduko ditugu. Tresna hauek 2. atalean finkatu diren helburu batzuk betetzen ez zituztenez sistema hauek baztertua izan dira.



### 4.1.1 Text to SNOMED

Sistema automatizatu bat da testu klinikoaren konbertsioa egiten duena (Patrick et. al, 2016). Hau da, testuan dagoen termino bat SNOMED CT-n agertzen bada, identifikatzailea eta dagokion kodea eta etiketa semantikoa (gaixotasuna, gorputz atala, ...) ematen duen sistema da. Sistema honek ingelesez dauden kontzeptu medikoak identifikatzen ditu osasun testuetan denbora errealean. Gainera, ikutu negatiboak dituzten eta kalifikatzaileak diren terminoak identifikatzeko gai da.

Sistema honen helburu nagusia testu hutsa hartzea eta SNOMED CT kontzeptuetan bihurtzea da. Horretarako, testua tokenizatu eta SNOMED CT kontzeptuen tokenizazioarekin lotu behar izango da. Ondoren testuan *chunk*-ak (edo testu zatiak) egingo dira beraien esanahian oinarrituz. Bukatzeko, sortu berri diren *chunk*-en gaitetik algoritmo bat aplikatuko (*TokenMatcher* deiturikoa) zaio ikutu negatibo (bi ezeztapen mota bereizten dira: aurre koordinatuak diren terminoak, hau da, termino negatiboa errepresentatzen duena SNOMED CT terminologian, eta esplizituki negatiboa direnak esaldi negatiboak dituelako) eta kalifikatzaileak identifikatzeko.

Beraz, erreminta hau baztertutzat jotzen dugu ingelesez dauden testuentzako soilik aplikagarria delako.

### 4.1.2 FreelingMed

Ixa taldeak Freeling analizatzailea medikuntzara egokitzearen ondorioz sortu zen FreelingMed (Maite Oronoz et. al, 2015). Sistema honen eginkizun nagusia gaztelaniazko osasun-txostenetan daude osasun-arloko terminoak (botikak, gaixotasunak, gorputz zatiak) identifikatzean datza.

Horretarako analizatzaile morfosintaktiko bat egokitu da SNOMED CT-ko (ikus 4.2 azpiatala) terminologiaren laguntzarekin. Gainera, etiketa sekuentziala garatu da eredu global bat jarraituz eta terminoen lista bat lortu, eta dagozkion CUI identifikatzaileak lortzen ditu.

Sistema honek bi hiztegi erabiltzen ditu. Lehenengoa hitz bakarreko terminoak soilik daude eta bigarrenean hitz anitzekoak. Bi hauek hainbat

laburtzapenekin eta SNOMED CT-rekin (4.2 azpiatala) aberasten dira. Espresio erregularren erabilera du tokenikazioan hainbat termino mediko termino bakarrentzat bezala identifikatzeko.

Aurretik aipatu den bezala, sistema hau baztertzen dugu gaztelaniaz dauden testuentzako aplikagarria delako.

Behin bi sistema hauek azalduta, SNOMED CT osasun terminologiari buruz hitz egingo da.

## 4.2 SNOMED CT : osasun terminologia

Azpiatal honetan, lehenik eta behin, irakurleari SNOMED CT zer den, zertarako den eta bere egitura nolakoa den azalduko dugu.

### Zer da?

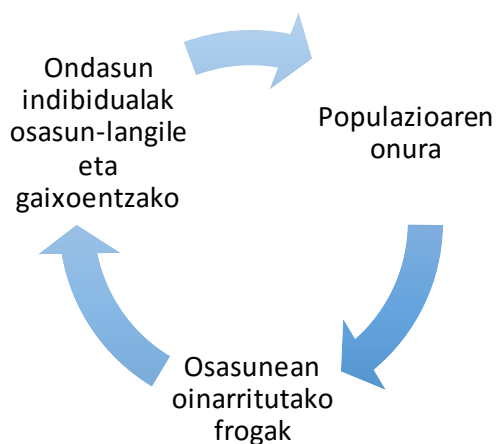
SNOMED CT (*Systematized Nomenclature of Medicine – Clinical Terms*) medikuntza alorreko azterketa zabala da. Beste hitzetan, osasun-txostenetan aurkitu daitezkeen terminoen deskribapenak, lotutako kontzeptuak edo erlazioak biltzen dituen ontologia da. Medikuntza-arloan orain arte sortu den terminologia eleanitz ulergarriena kontsideratzen da, eta hainbat hizkuntzetan eskaintzen da. Osasun langileei laguntzen dien erreminta da gaixoen informazioa erakusteko era egokian, zehatzean eta era nahastezinean aurkezteko.

Arrazoi hauengatik SNOMED CT-k gaur egun erabilera zabala du mundu osoan zehar. Hainbat hizkuntzatarako bertsioak daude edota garatze-lanetan, hala nola, holandesa, frantsesa, gaztelania,...

Erreminta honen jatorrizko hizkuntza ingelesa da, eta honentzako 300.000 kontzeptu baino gehiago definituak daude eta kontzeptu hauentzako 1.000.000 deskribapen inguru.

## Zertarako?

SNOMED CT informazio klinikoan oinarritzen denez, ondasun indibidualak ekartzen dizkie osasun-langileei eta gaixoei eta honekin batera populazioari. Gainera, osasunean oinarritutako frogak eguneratuak ekartzen ditu.



*Irudia 4. 1: SNOMED-CT-ren ondasun ziklikoa.*

## Egitura

Aurretik aipatu dugun moduan deskribapen anitz daude eta hauek kontzeptuei lotzen zaizkie erlazioen bidez, alde batetik definizio deskribatzaileak ditugularik (definizioa bera identifikatzeko balio dutenak), eta bestetik, definizioen sinonimoak daudelarik. Kontzeptu, deskribapen eta erlazio bakoitzak identifikazio zenbaki bakarra dauka (beraz, ez dira hauek errepikatzen) SNOMED *identifier* (hemendik aurrera identifikatzailea) izenekoa.

Hurrengo atalean SNOMED CT-ren egitura azalduko dugu. Horretarako, *International Health Terminology Standards Development Organisation* (IHTSDO) erakundeak sortutako SNOMED CT *browser*-ra erabiliko dugu errazago ulertzeko.

### 4.2.1 IHTSDO SNOMED CT Browser

Lehenik eta behin, *browser (IHTSDO SNOMED CT Browser, 2016)* hau hainbat hizkuntzatan aurki dezakegu (aurretik aipatutakoak, hau da, gaztelaniaz, ingelesez, suedieraz, nederlandera...). Guk ingelesezko bertsioa aukeratu dugu azalpenak egiteko.

#### Deskribapen motak

Kontzeptu bakoitzentzako bi deskribapen mota bereizten dira SNOMED CT-n: *Fully Specified Name (FSN)*, eta *Synonym* (bat baino gehiago izan ditzake kontzeptuko). Hau hobeto adierazteko, kontzeptu batekin (*Apoptosis*; Apoptosia euskaraz, 4.1 testu kutxa) azalduko dugu:

Apoptosis (morphologic abnormality)			
Term		Acceptability (US)	
F	Apoptosis (morphologic abnormality)	Preferred	(1)
S	Apoptosis	Preferred	(2)
S	Gene-directed cell death	Acceptable	(3)

Testu-kutxa 4. 1: "Apoptosis" kontzeptuaren deskribapen motak.

- *Fully Specified Name*: Kontzeptuak identifikatzeko eta bereizteko erabiltzen diren deskribapenak dira. Ez dira osasun-txostenetan aurkitzen diren terminoak. Termino honi parentesi artean kategoria semantikoa adierazten duen *semantic tag* deiturikoa zehazten zaio (kontzeptu honen gurasoa esan dezakegu dela). Aurreko taulan Apoptosis (morphologic abnormality) FSNa-ren kategoria semantikoa *morphologic abnormality* litzateke. Hau da, *Term* eremuak F erabiltzen du ((1) lerroa) FSN-ri erreferentzia egiteko. FSN bakarra edukiko dute kontzeptuek. FSN-ak direnak beti *Preferred* etiketa eramango dute.
- *Synonym*: Osasun-txostenetan dauden kontzeptuen deskribapenak dira hauek. Kontzeptu bakoitzerako gutxienez sinonimo bat zehazturik dago (3.1 testu kutxako (2) eta (3) lerroak) eta bi etiketen arabera sailkaturik daude: *Preferred* eta *Acceptable*. Termino hauek S letra erabiliko dute.

– *Preferred*: Hobetsitako sinonimoa adierazten da etiketa honen bitartez. Kontzeptu bakoitzerako hobetsitako sinonimo bakarra egongo da beti, hizkuntza edo dialekto bakoitzerako. Lehenengo testu kutxan Apoptosis litzateke S hizkia *Preferred* etiketa duelarik ( (2) lerroa).

– *Acceptable*: Gainerako sinonimo guztiak dira, hau da, hobetsitakoak ez diren sinonimoak. Kontzeptu bakoitzak hainbat sinonimo izan ditzake edo bakar bat ere ez. 3.1 testu kutxan, (3) zenbakia duena. Sinonimo honen izena Gene-directed cell death izango litzateke S hizkia eta *Acceptable* etiketa dituelarik.

## RF1, RF2 formatuak eta SNOMED CT-ko eta hierarkiak

SNOMED CT-ek bi formatu ditu, RF1 (*Released Format 1*) eta RF2 (*Released Format 2*), alegia. Gaur egun jadanik formatu bakarra erabiltzen da, RF2 (RF1 ordezkatur). Formatu hau bi zatitan banatzen da eta erabiltzaileen eskaerengatik sortua izan da. RF1 ordezkatzearen arrazoi nagusia RF2-k eskaintzen duen sendotasuna eta trinkotasuna da.

RF1 hiru fitxategitan banatzen da, hau da, fitxategiak kontzeptuetan, deskribapenetan eta erlaziotan sailkatzen dira. RF2-n, aldiz, fitxategi kopurua handitzen da (Erlazioak, Kontzeptuak, Deskribapenak, Crossmapak , ...) edukia mugatuagoa izanez.

SNOMED CT-k bi erro kontzeptu ditu. Lehenengoan metadatuak egituratzeko erabilitako erroa da: *Root Metadata Code*. Azken honetan, SNOMED CT beraren informazioa egituratzen da, *Fully Specified Name, Preferred, Acceptable, Current...* moduko elementuen bitartez.

Bigarren erroa, edukia antolatzeko erabiltzen da: *Root Concept Code*. Honen barruan 19 goi mailako hierarkia definitu dira (*Top Level Hierarchies*). Hierarkia hauetan kontzeptu guztiak barne bildurik daude, bakoitza dagokion hierarkian antolatuta. Kontzeptuen arteko ume-guraso loturak gauzatzeko, *is a* motako erlazioak erabiltzen dira. Ikusi 4.2 irudiko SNOMED CT-ko 19 hierarkiak.

- ▼ ● SNOMED CT Concept
  - ▶ ● Body structure (body structure)
  - ▶ ● Clinical finding (finding)
  - ▶ ● Environment or geographical location (environment / location)
  - ▶ ● Event (event)
  - ▶ ● Observable entity (observable entity)
  - ▶ ● Organism (organism)
  - ▶ ● Pharmaceutical / biologic product (product)
  - ▶ ● Physical force (physical force)
  - ▶ ● Physical object (physical object)
  - ▶ ● Procedure (procedure)
  - ▶ ● Qualifier value (qualifier value)
  - ▶ ● Record artifact (record artifact)
  - ▶ ● Situation with explicit context (situation)
  - ▶ ● SNOMED CT Model Component (metadata)
  - ▶ ● Social context (social concept)
  - ▶ ● Special concept (special concept)
  - ▶ ● Specimen (specimen)
  - ▶ ● Staging and scales (staging scale)
  - ▶ ● Substance (substance)

*Irudia 4. 2: SNOMED-CT-ren 19 hierarkiak.*

- ▼ ● Substance (substance)
  - ▼ ● Allergen class (substance)
    - ▼ ● Allergen or pseudoallergen (substance)
      - ▼ ● Drug allergen or pseudoallergen (substance)
        - ▼ ● Drug allergen (substance)
          - ▶ ● 3-Hydroxy-3-methylglutaryl-coenzyme A reductase inhibitor (substance)

*Irudia 4. 3: 3-Hydroxy-3-methylglutaryl...-en arbasoak.*

4.3 irudian substance kontzeptuaren gurasoa eta arbasoak ikus ditzakegu hierarkiaren bitartez deskribatuta. Gure kasuan sinonimoak euskaratzea beharrezko izango da GAP lan hau aurrera eraman ahal izateko.

Beraz, esan dezakegu kontzeptu bakoitzak *is a* erlazio bat izango duela beti. Hona hemen *Apoptosis* kontzeptuaren erlazioa (3.2 testu-kutxa):

Apoptosis (morphologic abnormality)			
Type	Description	Group	CharType
Is a (attribute)	Body structure, altered from its original anatomical structure (morphologic abnormality)	0	Inferred

Testu-kutxa 4. 2: : “Apoptosis” kontzeptuaren erlazioa (beste kasu batzuetan bat baino gehiago egon daitezke).

3.2 testu-kutxan ikus dezakegu nola *Apoptosis* terminoak erlazio bakarra duen eta erlazio mota hori *is a* den. Hau da, *Apoptosis*-en gurasoa beste kontzeptu batekin lotuko du *Body structure, altered from its original anatomical structure (morphological abnormality)* izango dena. Beraz:



Irudia 4. 4: Erlazioak.

SNOMED CT-n bestelako erlazioak ere badaude, hala nola, *causative agent, location, etab.*

Esan beharra dago honek CORENLP (4.3.1 azpiatala) ingeleseko analizatzaile linguistikoak SNOMED CT erabiltzen duela ingelesez dauden terminoen analisia egiteko.

## 4.3 Analizatzaile linguistikoak

Bi analizatzaile linguistiko erabili ditugu produktuan (CoreNLP ingeleseko terminoentzako eta Eustagger-lite euskarazko

terminoentzako). Biek testu hutsa jaso eta testu horren analisi linguistikoa bueltatuko dute.

### 4.3.1 CORENLP

#### Zer da?

Hizkuntza-analisia egiteko baliabideak eskaintzen ditu. Erakundeen izenak edo populazioa identifikatu, esaldien zati egokitan banatu, datak eta orduak normalizatu, POS-ak (*Part Of Speech*, kategoria gramatikala euskaraz) identifikatu, adierazi zein entitateari erreferentzia egiten dion izen-esaldi bakoitzak, klase-erlazioak lortu bi aipamen artean, etab. lortuko ditugu analizatzaile honen bidez.

Beraz, oso erabilgarria egingo zaigu *Stanford CoreNLP* izeneko analizatzaile multzoa testu analisi sakona egiteko. Analisisirako hurrengo hauek erabiliko ditugu:

- Lema
- Izen-entitateak
- Mendekotasun sintaktikoak
- POS (*Part Of Speech*)

*Stanford CoreNLP* Javaz idatzita dago (Java 1.8 zehazki) eta Linux, OS X eta Windows-en erabil daiteke.

#### 4.3.1.1 *Stanford CoreNLP* interfazea Python-entzako

`Med_en` eta `Med_eu` sistemak Python programazio-lengoaian inplementatuta daude (4.5 azpiatala). Hau dela eta, analizatzaile linguistikoak Python erabiliz lantzeko aukera izatea oso garrantzitsua da proiektu honentzako.

Python-entzako *wrapper*-a<sup>2</sup> da non Java-tik Python-era konbertsioa egiten duen (*Stanford CoreNLP – a suite of core NLP tolos, n.p*). Hau da, aurretik

---

<sup>2</sup> Programa bat, non programaren portaera moldatzen den honen funtzionamendua aldatu gabe.



esan dugun moduan *CoreNLP* Java-z idatzita dago eta *wrapper* honen bidez lortuko dugu zerbitzari honen kodea Python-ez bistaratzea.

### 4.3.2 *Eustagger-lite*

#### Zer da?

Euskararako etiketzaile/lematizatailea da non bere helburua hitz-forma bakoitzeko (bere testuingurua kontuan hartuta) dagokion lema eta hizkuntz etiketak ematea den (Ornoz, 2008). Analizatailea bi mailatan oinarrituta dago eta diseinatua izan da hiru modularekin: analizataile estandarra, aldaera linguistikoan analizatailea eta lemarik gabeko analisia.

Lema, POS (*Part Of Speech* edo kategoria gramatikala euskaraz) eta beste informazio morfoloikoa lortzeko gai da. Gainera, data, orduak eta zenbaki espresioak antzeman ditzake.

*Eustagger-lite* C-z idatzita dago eta web-zerbitzua eskaintzen du.

#### Anbiguotasuna

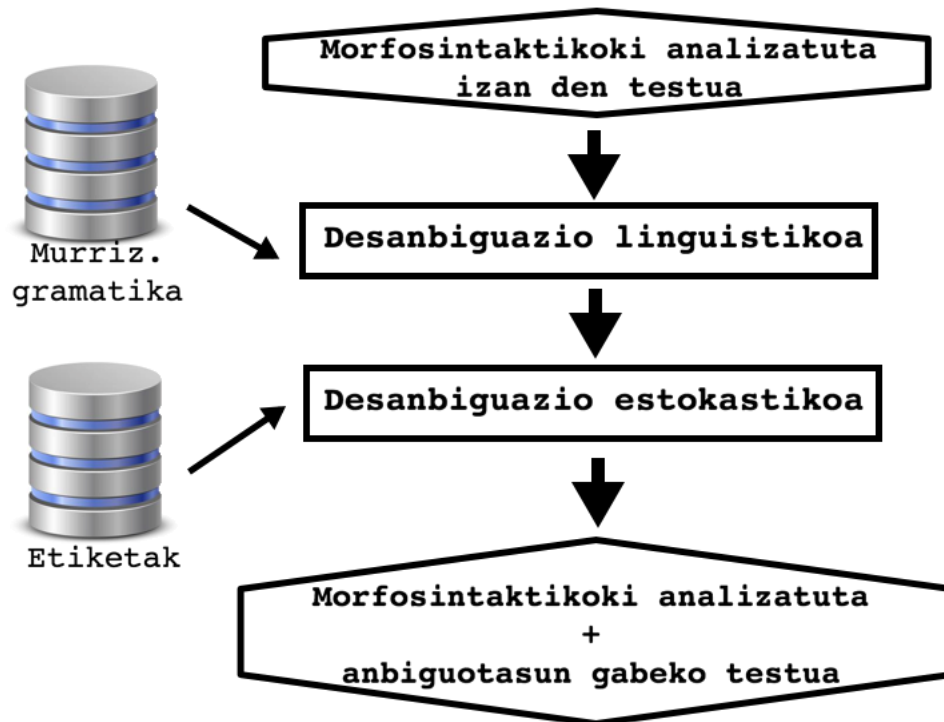
CG (*Constraint Grammar*) liburutegiko murriztapen erregelak erabiltzen ditu kategoria mailako anbiguotasuna txikitzeko.

Morfeus (IXA taldeak garatutako analizataile morfoloikoa) erraminta erabilia hitz-forma bati dagozkion analisi morfosintaktiko posible guztiak ematen ditu. *Eustagger-lite*-k aldiz, testuinguruko informazioa begiratzuz onartezinak diren interpretazioak baztertzen ditu. Honela, *Eustagger-lite*-k interpretazio bakarra eta egokia lortzen saiatuko da.

Hiru anbiguotasun morfosintaktiko mota aurki daitezke:

- *Kategoria*: izen/aditz, aditz/adjektibo/adberbio... modukoak non hitzak 2, 3 edo 4 interpretazio posible eduki ditzaketen. Bi teknika ezberdin konbinatzen ditu *Eustagger-lite*-k (hizkuntzan oinarrituan eta estatistikoetan) (ikusi 4.5 irudia).
- *Morfema ez-askeak*.

- *Sintaxia*: hitz bat subjektu edo objektu izan daitekeen.



*Irudia 4. 5: Eustagger-lite-ren arkitektura.*

Bukatzeko, estokastiko teknikak erabiliz etiketa bakarra aukera dezake (bat baino gehiago egon arren). Hau soilik erabilia, errore ratioa (hots, anbiguotasuna) %15-koa da, eta honi teknika linguistikoak gehitzen badizkiegu, errore ratioa %4,5-koa izango da.

#### 4.3.2.1 NAF formatua

Eustagger-lite-k emaitza hainbat formatuetan buelta dezake. Gure kasuan NAF formatuan izango da (Wordpress.let.vupr.nl., 2016). Hau, NLP (*Natural Language Processing*, informatika, adimen artifizial eta hizkuntzalaritzaren alorra) *pipeline* konplexuentzako diseinatuta izan da NLP diren arkitekturak erraz erabiltzeko. Hurrengo itxura hartuko du emaitza (4.3 testu-kutxa):

```

<?xml version='1.0' encoding='UTF-8'?>
<NAF xml:lang="eu" version="2.0">
<nafHeader>
  <linguisticProcessors layer="text">
    <lp name="EHU-eustagger" version="1.0.0"
beginTimestamp="2016-05-10T12:46:56+0200"
endTimestamp="2016-05-10T12:46:56+0200"/>
  </linguisticProcessors>
  <linguisticProcessors layer="terms">
    <lp name="EHU-eustagger" version="1.0.0"
beginTimestamp="2016-05-10T12:46:56+0200"
endTimestamp="2016-05-10T12:46:56+0200"/>
  </linguisticProcessors>
</nafHeader>
<text>
  <wf id="w1" offset="0" length="7" sent="1" para="1">Down-
en</wf>
  <wf id="w2" offset="8" length="9" sent="1"
para="1">sindromea</wf>
</text>
<terms>
  <!-- Down-en -->
  <term id="t1" lemma="down" morphofeat="NC000000" pos="N"
case="IZE ARR GEN NUMP MUGM ZERO HAS_MAI @&lt;IZLG
@IZLG&gt;">
    <span>
      <target id="w1"/>
    </span>
  </term>
  <!-- sindromea -->
  <term id="t2" lemma="sindrome" morphofeat="NC000000"
pos="N" case="IZE ARR BIZ- ABS NUMS MUGM @OBJ @PRED @SUBJ">
    <span>
      <target id="w2"/>
    </span>
  </term>
</terms>
</NAF>

```

Testu-kutxa 4. 3: "Down-en sindromea" terminoa NAF formatuan.

Beraz, eustagger\_lite eta NAF formatuaren bitartez hitz formaren informazio linguistikoa guztia lortuko dugu modu ulergarri eta erraz batean. Har dezagun , "Down-en sindromea" terminoa eta ikus dezagun zein informazio topa dezakegun NAF formatuan (4.3 testu-kutxa):

#### Eskainitako informazioa:

- Lemak: lemma="Down" eta lemma="sindrome"
- Anali linguistikoa (case etiketa):
  - o Hitzen arteko dependentziak: "@&lt;IZLG @IZLG&gt;" hizkiak -en atzikia adierazten dute. Kasu honetan 'Down-en'

- tokena 'sindromea' tokenaren izen laguna izen daitekeela( @ > IZLG) adierazten da (desanbiguatu gabe dago @IZLG > ere badugulako).
  - Kategoria (KAT) eta azpikategoria (AZP) nagusiak: IZE ARR (izen arrunta).
  - Kasu-markak (KAS): ABS (Absolutiboa).
  - Numeroa (NUM): NUMS (Numero singularra).
  - ...
- Hitzaren luzera: `offset="0" length="7"` eta `offset="8" length="9"`.
- POS (*Part Of Speech*): `pos="N"` eta `pos="N"` (biak izenak, *noun*).

## 4.4 Med\_en sistema

Sistema honetan oinarrituko gara `Med_en` euskarako bertsioa eraikitzeko. Sistema hau ingeleseko terminoetarako soilik baliagarria denez ez ditu euskararen terminoak egoki tratatuko (sintaxi ezberdina, morfologia, ...). Dena den, hurrengo orrietan erakutsiko dugu nola euskararako bertsioa sortzea dagoen (eta honekin batera, beste edozein hizkuntzarako) `Med_en` sistemaren kodea eskura izanez. `Med_en` sistemak *Stanford* analizatzailea erabiltzen du.

`Med_en` sistemak bi osagai nagusi ditu: `TermMed_en` eta `AnaMed_en`. Jarraian azalduko ditugu:

### 4.4.1 *termMed\_en*

Zerbitzari honetan kontzeptuak, deskribapenak, erlazioak, hizkuntza eta hierarkiak jasoko ditu (kargatu). Zerbitzari hau martxan jarri ahal izateko *CoreNLP (Stanford)* zerbitzaria martxan eduki behar izango dugu deskribapeneri dagozkien informazioa lortzeko (lema, POS,...).

`TermMed_en` zerbitzariaren bitartez *SNOMED-CT browser*-ak erakusten duen hainbat informazioa lortuko dugu (FSN, SNOMED ID-ak, hierarkiak etab.).

Behin hau prest egonda, `anaMed_en` termino analizatzailea erabili ahal izango dugu ingeleseko terminoetatik informazio baliagarri guztia lortzeko.

#### 4.4.2 anaMed\_en

Esan dugun bezala, bezero honek funtzionamendu egoki bat izateko termMed\_en zerbitzaria martxan eduki behar dugu (baliabide guztiak kargatuak egoteko).

Analizatzaile honetan modu egokian tratatu izan ez den informazioaren (eponimoak identifikatu lema modu egokian agertzeko, lemaren arabera bilaketak egin etab.) zuzenketa egin da. Ikusi dugun bezala eponimoek (izen bereziak dituzten gaixotasunak, adibidez) berebiziko garrantzia izango dute terminoen analisia egiteko. Gainera, MedTerm\_en-eko SNOMED identifikatzailea eta Hierarkia esleituko zaio terminoari, AnaMed\_en bezeroaren bitartez.

## 4.5 Python eta Django-a

Produktua egiteko lana bi zati nagusitan banatua izan da: lehengo zatia Med\_eu sistema (produktu zati handiaren inplementazioa hemen egin da) eta bigarren zatia Django *framework*-aren inguruan (Med\_eu sistemaren beharra izango duen interfazea). Bi hauek erabiltzea erabaki oinarrian hartu diren tresnak hauek erabiltzen dituztelako (Med\_en sistema Pythonez zegoen eta aurretik eginda zegoen interfazea Django-n *framework*-a erabiltzen zuen).

### 4.5.1 Programazio-lengoaia

Med\_eu sistema egiteko erabili den lengoaia Python izan da. Programazio-lengoaia honen ardatz nagusia sintaxi garbi, erraz eta irakurgarria da (tabulazio egokia erabiltzea, adibidez).

Paradigma anitzeko lengoaia da, objektuetara zuzenduriko programazioa, programazio inperatiboa eta, neurri txiki batean, programazio funtzionala jasaten ditu. Lengoaia interpretatua da eta plataforma anitzetan erabiltzen da.

Lengoaia hau erabili da hainbat arrazoiengatik:

- Med\_en sistema Pythonez dagoelako.
- Kodea ulertzeko erraza delako.

- Kode lerro gutxi batzuk erabiliz funtzio indartsuak sortu ahal dira (beste programatzaileek aurretik sortu dituzten paketei (`import`) esker).
- Edozein sistema eragilean erabili daiteke.
- Ondoren sortu behar den interfazea Django-n *framework*-ean egin ahal izateko.

#### 4.5.2 Django

*Framework* bat da kode irekikoa Pythonez idatzita dagoena. MVC (*Model Controller View*, eredu-bista-kontrolatzailea euskaraz) arkitektura erabiltzen du.

*Framework*-ak hurrengo itxura hartuko du proiektu bat sortzerakoan:

```

/nireProiektua/
  manage.py
  nireProiektua/
    __init__.py
    settings.py
    urls.py
    wsgi.py

```

Testu-kutxa 4. 4: Django-ren hasierako fitxategiak.

Fitxategi hauek hurrengo hau adierazten dute:

`/nireProiektua/`: direktorio erroa. Errepositorio hutsa.

`manage.py`: proiektuarekin elkarrekintza baimentzen du.

`nireProiektua/`: proiektuaren Python paketea.

`nireProiektua/__init__.py`: fitxategi hutsa da non Python paketearen direktorioa bihurtzen duen.

`nireProiektua/settings.py`: Django proiektua konfiguratzeke ezaugarriak definitzen dira.

`nireProiektua/urls.py`: proiektuaren URL erazagupenak agertzen dira. Era batean esanda, proiektuaren aurkibidea izango litzateke.

`nireProiektua/wsgi.py`: zerbitzaria atzitzeko ezaugarriak definitzen dira.

`nireproiektua/views.py`: ekintzak sortzeko funtzioak hemen implementatuko dira. Fitxategi hau gizakiak sortu beharko du.

nireproiektua-ren direktorioaren barnean hainbat direktorio eta fitxategi gehiago aurki daitezke (proiektuaren beharretara moldatuz): `.html`-ak, `.css`-ak, `.js`-ak, ...

## INTERFAZEA SORTZEKO BEHARREZKO TEKNOLOGIAK:

Medikuntzarako terminoen informazioa lortzeko, interfaze bat garatu da. Interfaze honek informazioa modu erakargarrian ikusteko, web-orriak sortzeko eta diseinatzeko hainbat teknologia mota erabili ditugu.

### HTML:

HTML (*HyperText Markup Language*) markaketa lengoaia bat da, web-orriak egiteko sortu zena. Informazioa egituratzeko erabiltzen da — testua burualdean, paragrafo batean, zerrenda batean eta beste zenbaitetan sailkatuz — eta neurri batean itxura eta semantika ere deskribatzeko erabil daiteke.

HTML5 (HTMLko 5. bertsioa) *World Wide Web*-en (elkarloturiko hipertestuko dokumentuen sistema) oinarritzko hizkuntza den HTML-ren bosgarren berrikuste garrantzitsua da.

Informazio guztia HTML etiketen tartean egon behar da. Etiketak, `<etiketa>` `</etiketa>` edo `</etiketa>` itxura izango dute beti.

Hona eredu simple bat:

#### HTML kodea

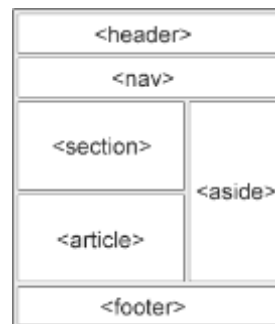
```
<!DOCTYPE html>
<html>
  <head>
    <title>Page
Title</title>
  </head>
  <body>
    <h2>Hau
goiburukoa da</h2>
    <p>Hau testua da</p>
  </body>
</html>
```

#### Nabigatzailean bistaratuko dena

**Hau goiburukoa da**  
Hau testua da

Taula 4. 1: HTML kodearen eredu bat.

HTML5-ek elementu semantiko berriak eskaintzen ditu web-orri baten atal ezberdinak identifikatzeko (3.6 irudia).



Irudia 4. 6: HTML5-en elementu semantikoak.

Behin web-orria eginda, diseinuaren beharra izango dugu erakargarriagoa izateko. Honetarako CSS programazio-lengoaian definituriko estilo orriak beharko ditugu.

## CSS:

CSS (*Cascading Style Sheets* edo euskaraz kaskadako estilo-orriak) HTML edo XML (beraz, XHTML ere bai) lengoaian egituratutako dokumentu baten aurkezpena definitzeko balio duen lengoia da. W3C<sup>3</sup> da beronen definizioa mantentzen duen erakundea.

CSS lengoiaren helburu nagusia dokumentu baten edukia eta aurkezpena bereiztea da. CSS-ak hiru modu desberdinetan erabili daitezke:

### HTML + CSS

```
<!DOCTYPE html>
<html>
  <head>
    <title >Page Title</title>
    (3) <link rel="stylesheet"
    type="text/css"
    href=".cssFitxategiarairistekoPath" />
    (2) <style>
      p {
        font-family: "Verdana";
        font-size: 20px;
      }
    </style>
  </head>
  <body>
    (1) <h2 style="color: red" >Hau
    goiburukoa da</h2>
    <p>Hau testua da</p>
  </body>
</html>
```

Nabigatzailean bistaratuko dena

**Hau goiburukoa  
da**  
Hau testua da

Taula 4. 2: HTML eta CSS-aren eredu bat

<sup>3</sup> World Wide Web Consortium: *World Wide Web-erako* estandarrak sortzen dituen erakundea da



- **Etiketa bertan erabiltzea:** honela egin nahi badugu etiketa batean barruan `style=""` hitza erabili beharko dugu (aurreko taulako (1)).
- `<head>` etiketaren barruan `<style>` etiketa erabiltzea. (aurreko taulako (2)). Kasu honetan, `<p>` etiketa duten elementuei CSS aplikatuko zaie (verdana letra-estiloa eta 20px-eko tamaina ezarriz).
- Erabiliena da: Beste toki batean dagoen `.css` fitxategiari erreferentzia egiten dio.(aurreko taulako (3)).

### **id eta class atributuak CSS munduan:**

Bi atributu hauek etiketen artean idazten dira. Atributu hauen bidez etiketak identifikatzeko erabiltzen dira propietateak errazago esleitu ahal izateko (letren koloreak, letren tamaina, posizioak, ...). Normalean aurreko taulako (3) moduarekin konbinatzen da. `id` eta `class`, atributuen arteko diferentzia hurrengo hau da:

- `id`: errepikaezinak dira, hau da, bakarra. (“#” batez identifikatzen dira CSS estilo orrian).
- `class`: errepikatu daitezke. Hau da, etiketa desberdina `class` atributu berdina eduki dezakete (“.” batez identifikatzen dira CSS estilo orrian).

`id` eta `class` hauek dagokien estilo orrian bilatuko dute bere atributu baliokidea eta dagozkien propietateak esleituko zaie etiketei. Hau egiterakoan, etiketek itxura bat izango dute. Hona hemen hobeto ulertzeko adibidea (4.5 testu-kutxa):

`stylesheet.css` estilo orrian:

```
#kolorea { /*id mota*/
    color: red; /* letra gorria */
}
.lettraMota { /*class mota*/
    color: #b00ff0; /* letra lila */
}
```

*Testu-kutxa 4. 5: CSSko id eta class identifikatzaileak.*

HTML + CSS	Nabigatzailean bistaratuko dena
<pre> &lt;!DOCTYPE html&gt; &lt;html&gt;   &lt;head&gt;     &lt;title &gt;Page Title&lt;/title&gt;     &lt;link rel="stylesheet" type="text/css" href=" stylesheet.css " /&gt;   &lt;/head&gt;   &lt;body&gt;     &lt;h2 id="kolorea" &gt;Hau goiburukoa da&lt;/h2&gt;     &lt;p class="letraMota"&gt;Hau&lt;/p&gt;     &lt;p class="letraMota"&gt;testua&lt;/p&gt;     &lt;p class="letraMota"&gt;da&lt;/p&gt;    &lt;/body&gt; &lt;/html&gt; </pre>	<p>Hau goiburukoa da Hau testua da</p>

Taula 4. 3: HTML eta CSS id eta class atributuekin eredu bat

Beraz, web-orriari diseinua ezarrita, orain dinamikoki kargatzeko funtzioak gehitu behar zaizkio.

## JAVASCRIPT:

*JavaScript* (edo JS) interpretatutako programazio-lengoaia bat da, Java (objektuei zuzendutako programazio-lengoaia) programazio-lengoiarekin antzekotasun handia duena, baina web-orrien garapenera zuzenduta dagoena.

Hasiera batean HTML kodean txertatuta erabiltzen hasi zen, bezero aldeko lanak egiteko. Baina Web 2.0-aren iritsierarekin batera, AJAX edo JCC bezalako teknologiak erabiltzea ere ahalbidetzen du.

Goazen adibide batekin hobeto azaltzera. Adibide hauek hurrengo hau egingo du: <p> etiketaren ganean klik egiterakoan, klik egin deneko data eta ordua azalduko dira <div> etiketa batean. Horretarako `onclick()` deritzon ebentoa eta non nahi dugun datuak azaltzea esan beharko diogu. Kasu honetan *demo* id-a duen <p> etiketan data eta ordua azalduko da.

HTML+ JS	Itxura klikatu aurretik	Itxura klikatu ondoren
<pre> &lt;!DOCTYPE html&gt; &lt;html&gt; &lt;body&gt;  &lt;p onclick="document.getElementById Id('lor্তু').innerHTML = Date()"&gt; Lortu data eta ordua&lt;/p&gt; &lt;p id="lor্তু"&gt;&lt;/p&gt;  &lt;/body&gt; &lt;/html&gt; </pre>	Lortu data eta ordua	Lortu data eta ordua  Sat Jun 11 2016 18:17:15 GMT+0200 (CEST)

Taula 4. 4: JS eta HTML-ren eredu bat

Bukatzeko, AJAX teknologia erabilia izan da JavaScript hainbat funtzioetan (ikusi 6.36, 6.38, ... testu-kutxak) AnaMed\_eu bezeroa eta sct4text funtzioa deitzeko (6.6 Azpitala).

## AJAX:

Ajax (*Asynchronous JavaScript and XML*) hurrengo teknologiak bateratuz web elkarrenragileen garapenerako teknika da. Zerbitzariarekin informazio gutxi trukatzuz web-orrialdeak harkorrangoak bihurtzeko asmoarekin sortu zen. Modu horretan ez dago orrialdea errefreskatu beharrik erabiltzailea eskatzen duenean. Web-orrialdeen abiadura, elkarreragina, funtzionalitatea eta erabilgarritasuna handitzen ditu.

Ajax teknologiari esker, erabiltzaileak ez du zuzenean web-zerbitzarira eskaera egiten; Ajax motoreari eskaera egin, eta berau arduratzen da web-zerbitzariarekin konexioa egiteaz. Motorearen bitartez informazioa jasotzean, bistaratu egiten da. Guzti honi esker AJAX aplikazio batek mahaigaineko aplikazio batek bezalaxe lan egin dezake, non datuak eguneratzeko pantaila osoa eguneratu beharrik ez dagoen.

## HTML + JS + AJAX

```
<!DOCTYPE html>
<html>
<body>
<div id="demo"><h3>AJAX testu hau
aldatuko du</h3></div>

<p      onclick="loadDoc()">Aldatu
formatua</p>

<script>
function loadDoc() {
    var      xhttp      =      new
XMLHttpRequest();
    xhttp.onreadystatechange      =
function() {
        if(xhttp.readyState==4&&
xhttp.status      ==200) {
document.getElementById("demo").i
nnerHTML = xhttp.responseText;
        }
    };

    xhttp.open("GET", "ajax_info.txt",
true);
    xhttp.send();
}
</script>
</body>
</html>
```

Itxura klikatu Itxura klikatu ondoren aurretik

**AJAX testu hau aldatuko du**

Aldatu formatua

AJAX ez da programazio lengoai berria

AJAX web-orri dinamikoak eta azkarrak sortzeko teknika bat da.

Taula 4. 5: AJAX, JS eta HTML-ren eredu bat

Adibide honekin ikusi dezakegu nola `<p>` etiketa testuaren gainean klik egiterakoan zegoen testua desagertuko (`<div>` etiketan agertzen dena kenduz ere) den, testu berria agertuz (zerbitzariari dei eginez `xhttp.open` funtzioaren bitartez).

Bukatzeko esan beharra dago, AJAX asinkronoa dela. Beraz, honen bidez web-orri azkarrak sortu ditzake. Modu sinkronoan erabili daiteke AJAX ere.

Erabiliko dugun teknologia azaldu ondoren, hurrengo orrietan (5. Atala) `Med_en` eta `Med_eu` sistemen analisia egingo dugu eta haien klase-diagramak eta erabilpen-kasuak azalduz.

# 5

---

## Analisisa

Esan dugun bezala, lehendabizi Med\_en sistemaren analisisa egingo dugu. Atal honetan argi ikusiko da zer zegoen egina eta zer aldatu den Med\_eu sistema sortzeko. Ondoren, Med\_eu sistema erabiliko duen aplikazioari buruz hitz egingo dugu, OsatEUS deiturikoa.

### 5.1 Arkitektura

Med\_en sistema hiru zatitan banatzen da (bi zerbitzari eta bezero bat):

**Zerbitzariak:**

- CoreNLP (corenlp.py): *Stanford* unibertsitateko analizatzaile linguistikoa da zerbitzari moduan erabilia. Terminoen analisisa ingelesez

egiteko baliabideak eskaintzen ditu. Terminoen informazio lexikoa (lema edo *POS*, adibidez) zerbitzari honetatik lortuko dugu.

- `TermMed_en` (`ts_en.py`): baliabideak kargatuko dira (kontzeptuak, terminoak, deskribapenak, ...) hainbat prozesu igaro ondoren. Baliabide hauek era egokian kargatzeko, `klaseak.py` moduluaren beharra izango dugu, non *concept* edo *description* bezalako ezaugarriak hiztegi (*dictionary*) moduan kargatuko diren<sup>4</sup>. Zerbitzari honetatik terminoen FSN, hierarkiak edo ID-ak lortuko dira, esate baterako.

### **Bezeroa:**

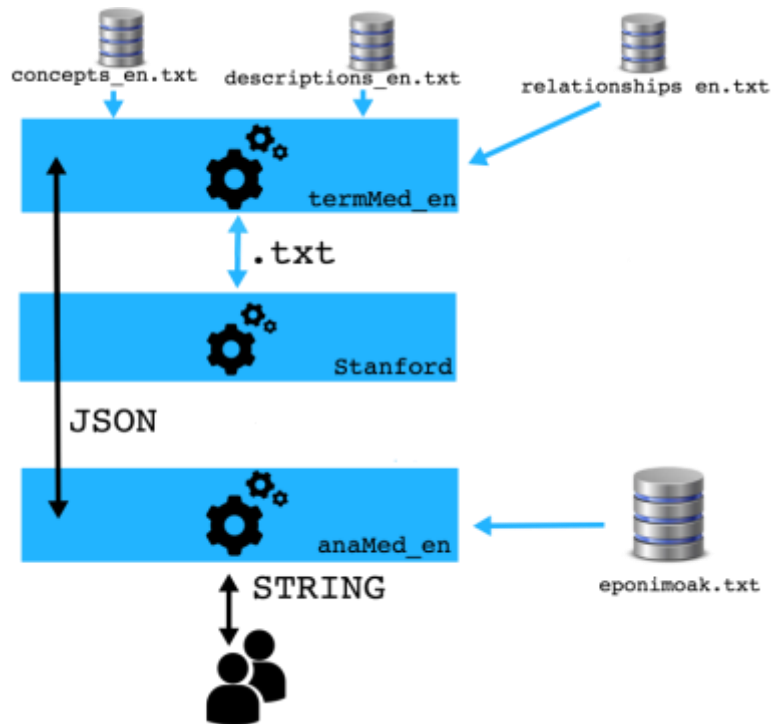
- `AnaMed_en` (`analizatzailea_en.py`): eponimoak eta antzeko hitz bereziak modu egoki batean tratatuko ditu. Horretarako, `eponimoak.txt` fitxategia beharko du (eponimo guztiak hemen aurkituko dira). Ingelesko terminoaren analisia bueltatuko dio erabiltzaileari (hierarkia, *sctId*, ...). Esan den bezala, `TermMed_en` zerbitzariari deituko dio terminoaren hierarkiak edo ID-ak terminoaren analisiari gehitzeko.

Bukatzeko, esan beharra dago 5.1, 5.2 eta 5.3 irudietan agertzen diren `.txt` guztiak Olatz Perez de Viñaspre-renak direla. Hona hemen `Med_en` arkitectura moldaketak egin aurretik eta ondoren (hainbat aldaketak egin dira sistema hobetzeko asmoarekin).

---

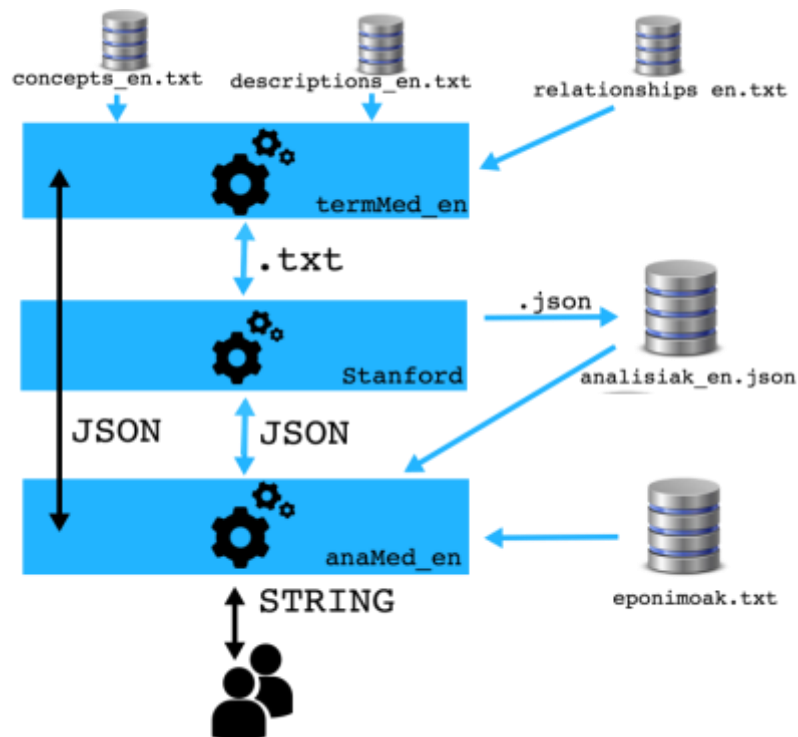
<sup>4</sup> Beste hizkuntzetan (Javan adibidez) *hash* izena edukiko zuten, hau da, *dictionary* = *hash*

**MED\_EN SISTEMAREN ARKITEKTURA AURRETIK:**



*Irudia 5. 1: Med\_en sistemaren hasierako arkitektura.*

**MED\_EN SISTEMAREN ARKITEKTURA ONDOREN:**



*Irudia 5. 2: Med\_en sistemaren bukaerako arkitektura.*

Ikusten den bezala aldaketa esanguratsuen `analisiak_en.json` (ikusi 6.3.1.1 azpiatala) fitxategiaren sorkuntza izan da. Fitxategi honen barruan termino bakoitzaren informazio lexiko guztia gordeko da `TermMed_en` kargaketa askoz azkarragoa izateko (ikusi 6.3 azpiatala). Beraz, `AnaMed_en` soilik *Stanford*-en analisia beharko du `analisiak_en.json` fitxategitik edo termino batzuk ezabatuak izan direnean hauen kargaketa egiteko.

`Med_en` sistema bi zatitan banatuko da ere (kasu honetan, bezero eta zerbitzari bat edukiz):

### Zerbitzaria:

- `termMed_en` (`ts_eu.py`): baliabideak kargatuko dira (kontzeptuak, terminoak, deskribapenak,...) hauen analisia egin ondoren. Baliabide hauek era egokian kargatzeko, `klaseak_eu.py` moduluaren beharra izango dugu. Kasu honetan ez dira hiztegietan gordeko, baizik eta objektuetan graduan zehar objektuekin lan egin dugulako. Horretarako, lau klase berri sortu ditugu gure `klaseak_eu.py`-n:

- o *concept*: hemen kontzeptuak gordeko dira.
- o *relationship*: kontzeptuen arteko erlazioak adieraziko dituzte.
- o *description*: kontzeptuen deskribapenak biltegitatuko dira.
- o *language*: deskribapena zein hizkuntzan egongo den adieraziko du.

Lau objektu hauen atributu guztiak klase-diagraman daude (4. 4 irudia). *Stanford* zerbitzaria *wrapper* (ikusi 6.4 azpiatala) batekin ordezkatzeko dugu `eustagger-lite` erreminta erabiltzen duena:

- o *Wrapperra* (`wrapper_eustagger.py`): euskarazko terminoen analisia egingo da. Horretarako `Eustagger-lite` erreminta erabiliko dugu. Behin `Eustagger-lite`-k analisia eginda NAF formatutik JSON formatura pasako da terminoa.

Ingeleseko bertsioan bezala, zerbitzari honetatik erabiltzaileak sartu dituen terminoaren FSN, hierarkia edo SNOMED ID-ak lortuko ditu, besteak beste.

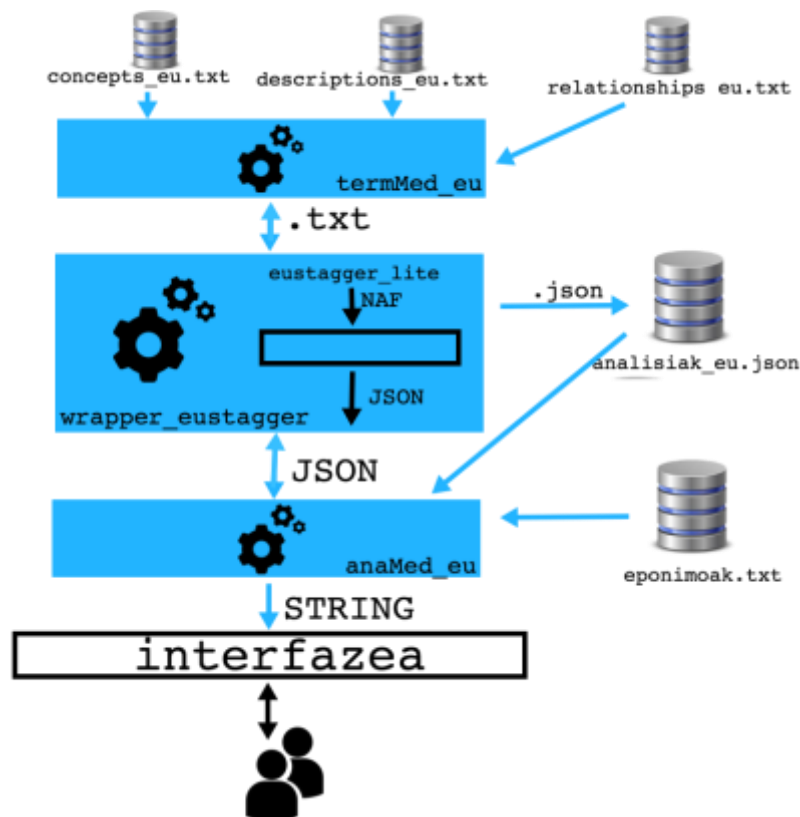
### Bezeroa:

- `anaMed_en` (`analizatzailea_eu.py`): ingeleseko funtzionalitate berdina edukiko du. Hemen, eponimoak modu sakonago batean tratatu behar izango dira euskarak duen morfosintaxiarengatik



(Down eponimoa lortu behar izango dugu Down, Down-en eta Downen hitzekin). Ingelesko bezeroan bezala, termMed\_eu zerbitzariari deituko dio erabiltzaileak idatzi duen terminoaren hierarkiak edo ID-ak analisiari gehitzeko.

Interfazean, erabiltzaileak idazten duenaren (testu hutsa) analisia egin eta terminoak detektatuko ditu pantailaren eskuinaldean adieraziz zeintzuk diren etiketa kolore mota desberdinak erabiliz (hierarkiaren arabera koloreak esleituko dira). Hau egiteko AnaMed\_eu behar izango da:



Irudia 5.3: Med\_eu sistemaren bukaerako arkitektura.

Adibide baten bidez hobeto azalduko dugu. Horretarako, eponimo bat duen termino bat aukeratu dugu “Downen sindromea” izenekoa:

**AURREBALDINTZA:** TermMed\_eu zerbitzaria martxan egotea.

**URRATSAK:**

- 1.- Erabiltzaileak interfazeko testu-kutxa batean Downen sindromea idaztea eta ondoren kurtsorea testu-kutxatik kentzea.
  - 2.- Interfazeak AnaMed\_eu deituko du.
  - 3.- AnaMed\_eu konprobatuko du ea termino bat den.
- Horretarako, analisiak\_eu.json fitxategian begiratu beharko du ea

Downen sindromea dagoen. Honengatik, TermMed\_eu zerbitzaria martxan egon behar du.

4.- Hau baldinba dago fitxategian Downen sindrome-aren informazio lexikoa, SNOMED Identifikatzailea eta hierarkia bueltatuko du JSON formatuan.

5.- AnaMed\_eu konprobatuko du ea eponimorik dagoen terminoan eta hau modu egokian tratatuko du (Downen-etik Down-ra pasako da izatera). Gainera, termino luzeenaren kontrola egingo da (termino anitza).

6.- Terminoa denez, interfazearen eskuinaldean kolore batez terminoa etiketatua agertuko da (ikusi 6.6 azpiatala). Kasu honetan, etiketa gorrixka izango da.

Behin Med\_en eta Med\_eu sistemen arkitekturak azaldu ondoren, hauek programazioari dagokionez nola dauden antolatuta ikusiko dugu, klase-diagramak azalduz. Lehendik, Med\_en klase-diagramarekin hasiko gara.

## 5.2 Klase-diagrama

Med\_en klase diagramaren azalpena (kolore laranja):

Med\_en klase diagrama lau kudeaketa-kontzeptuetan banatuta egongo da (*DescriptionList*, *LanguageList*, *RelationshipList* eta *ConceptList*) non hiztegietan gordeko diren (beste lengoaietako *hash*-ak) deskribapenak, hizkuntzak, erlazioak eta kontzeptuak,.

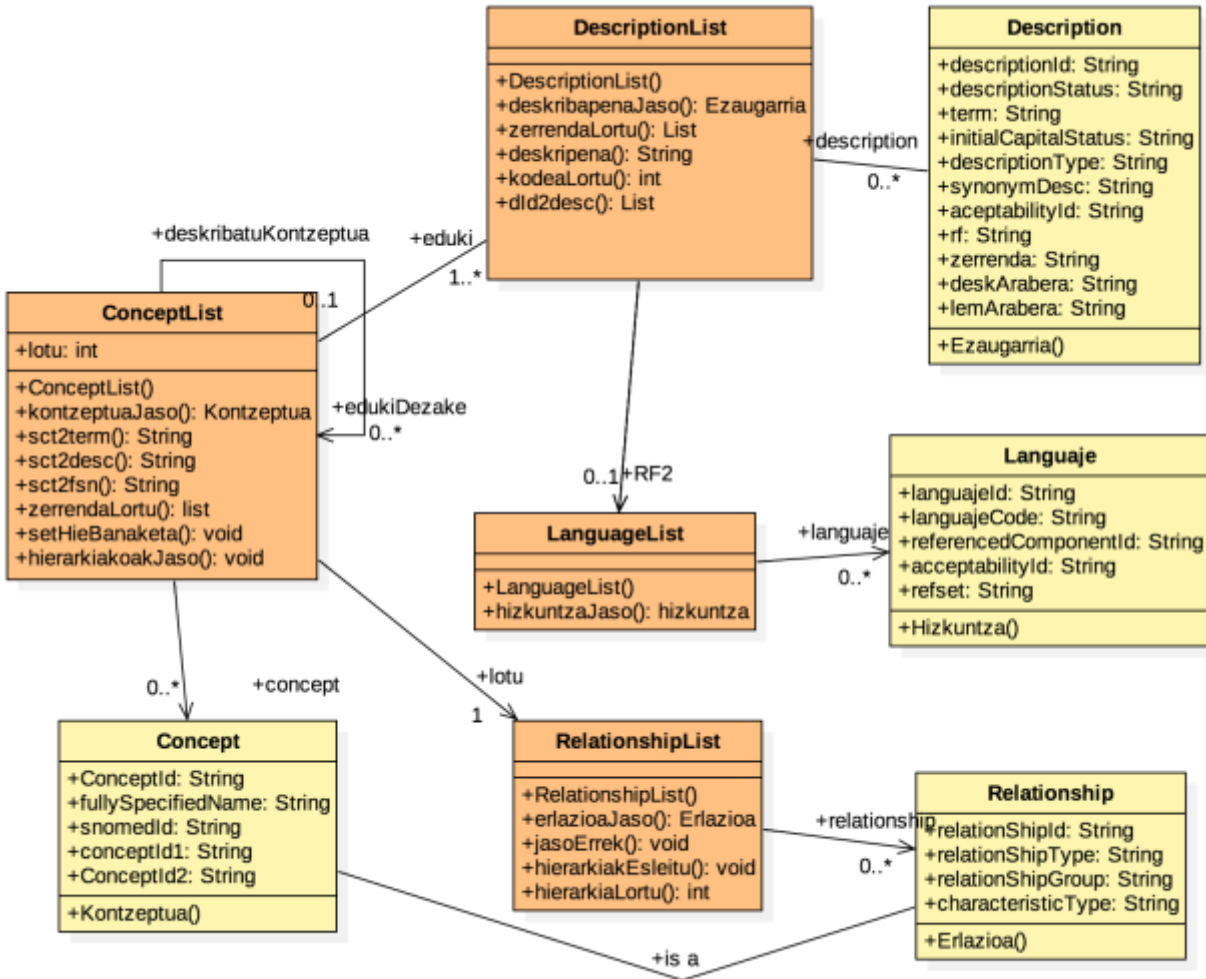
*ConceptList*-etan kontzeptuak gordeko dira eta berdina beste hiru kudeaketa-kontzeptuekin.

Kontzeptu batek beste kontzeptu batzuk eduki ditzake eta gainera, *ConceptList* batek *DescriptionList* bat edo gehiago edukiko ditu (*FSN* edo *Synonym* bezalakoak). *DescriptionList* hizkuntz (*LanguageList*) bakarrean edo bitan egon daiteke bere formatuaren arabera (RF1 edo RF2). Bukatzeko, *ConceptList* batek *RelationshipList* bat izango du kontzeptuen arteko erlazioak ezartzeko (*isa* motakoa kontzeptu umeak eta gurasoak ezberdintzeko).

Med\_eu sistemarentzako lau objektu-kontzeptu sortu izan dira (*Concept*, *Description*, *Language* eta *Relationship*) eta hauek Med\_en-en zeuden kudeaketa-kontzeptuekin lotuko dira. Lau objektu hauek klase-

diagraman horiz agertuko dira eta ez dute Men\_en-en funtzionamendua aldatuko. Objektuko atributu bakoitzari dagozkien get() eta set() funtzioak ez dira azalduko, klase-diagrama handiegia ez izateko.

Hurrengo itxura hartuko du klase diagramak:



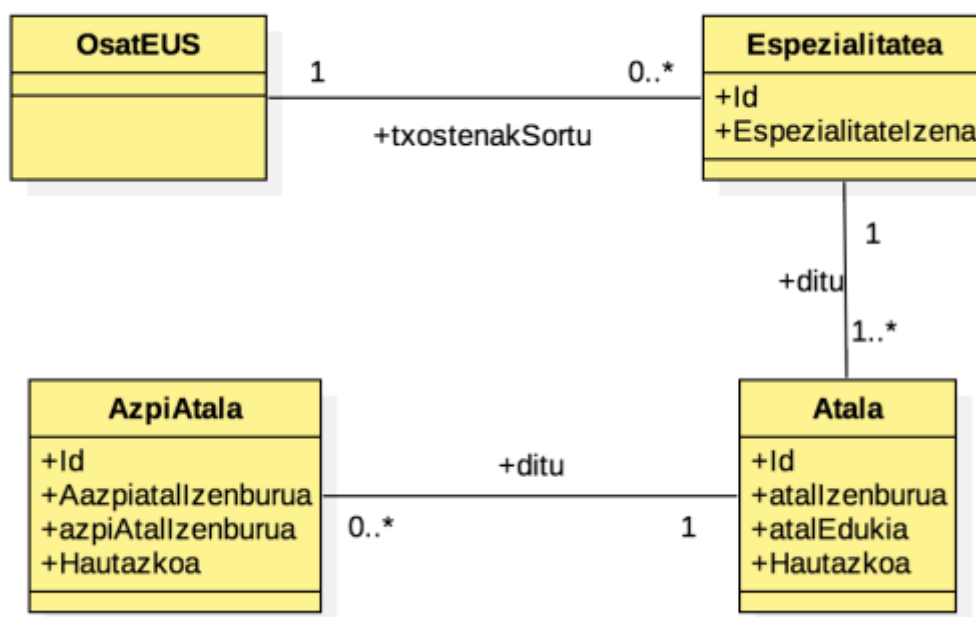
Irudia 5. 4: Bi sistemen klase-diagrama.

Bi sistemen klase-diagramak erakutsi ondoren, OsatEUS aplikazioaren egiturari buruz hitz egingo dugu aplikazio honek duen elementuak erakutsiz.

## 5.3 OsatEUS egitura

Aurretik erdi eginda zegoen osatEUS aplikazioari buruz baliatu gara. Egin den aldaketarik nabarmentsuena espezialitate bakoitzaren interfazea moldatzea izan da, erabiltzaileak sartzen dituen terminoak detektatzeko.

Aplikazio honen alta-txosten (Basabe, 2014) eremuan (ikusi E eranskina) espezialitate, atal eta azpiatalez osatuta dago. Hau da, aplikazioaren datu-basean hainbat espezialitate mota daude (aparatu-digestioa, ...). Espezialitate bakoitzak hainbat atal edukiko ditu (gutxi batzuk hautazkoak izango direnak) eta atalak azpiatalak eduki ditzakete. Atal eta azpiatal eremuetan erabiltzaileak osasun-txostenari dagokiona idatziko du. Osasun-txostenak era egokian bete ondoren, datu-basean gordeko dira identifikatzaile bat emanez erreferentziatzat. Hauxe da bere egitura (5. 5 irudia):



Irudia 5. 5: OsatEUS aplikazioaren egitura.

Hona hemen hobeto ulertzeko alta-txosten bat (alta-txosten honetan ez da testu guztia agertuko kasu honetan ez delako garrantzitsua):

ALTAREN TXOSTEN KLINIKOA (41)	
<b>ANESTESIOLOGIA ETA BERPIZTEA</b>	
<p><b>OSPITALERATZEKO/KONTSULTARAKO ARRAZOIA.</b> Tenporalaren eta <b>masailezurraren</b> arteko artroplastia egin ondoren, lekualdaturiko 67 urteko gizonezkoa.</p> <p><b>ALTAREN ARRAZOIA.</b> Aurpegi-masailetakoa eta ahoko Kirurgiako Zerbitzura lekualdatzea.</p> <p><b>AURREKARIAK.</b> Alergia die penizilinari, pobidona iododunari eta <i>Anisakisari</i>; erretzaile ohia da, hipertentsio arteriala du eta ...</p> <p><b>EGUNGO GAIXOTASUNA.</b> Aspaldiko traumatismoaren ondorioz, pazienteak <b>masailezurraren</b> eta tenporalaren arteko ankilosia zuen. 2013ko azaroaren 27an, programatutako ebakuntza egin zaio...</p> <p><b>EBOLUZIOA</b> <b>BIZKORKETARAKO GELAN.</b> Kontziente dago, arnas autonomia du eta hemodinamikoki egonkor dago.</p> <p><b>ALTA EMATERAKOAN.</b> Tentsio arteriala, 100/50 mm Hg. Tratamendu basoaktiborik gabe. Bihotz-maiztasuna, 54 tau/min; arnas... <b>BIHOTZAREN AUSKULTAZIOA.</b> Murmurio erritmikoak. <b>BIRIKEN AUSKULTAZIOA.</b> Arnas murmurioa entzuten da... <b>SABELALDEA.</b> Biguna eta zanpagarria; haztatzean... <b>GILTZURRUNAK.</b> Diuresi egokia, medikaziorik gabe.</p> <p><b>PROBA OSAGARRIAK.</b> <b>BIOKIMIKA.</b> Kreatinina, 1,41 mg/dl. <b>HEMOGRAMA.</b> Hemoglobina, 12,3 g/dl; plaketak, 280.000; leukozitoak, 8.300. <b>KOAGULAZIOA.</b> Nazioarteko ratio normalizatua (INR), 1,1.</p> <p><b>DIAGNOSTIKOA TENPORALAREN ETA MATRAILEZURRAREN ...</b></p> <p><b>TRATAMENDUA ETA GOMENDIOAK.</b> Anestesiako Zerbitzuaren aldetik ez du tratamendu edo jarraipen berezirik behar. Tratamendua, Ahoko eta...</p>	
	Sinadura-data1/1 orrialdea

Hona hemen eremu bakoitzaren sailkapena:

- **ESPEZIALITATEA (beltza):** ANESTESIOLOGIA ETA BERPIZTEA.
- **ATALAK (urdina):** OSPITALERATZEKO/KONTSULTARAKO ARRAZOIA, ALTAREN ARRAZOIA, AURREKARIAK, EGUNGO GAIXOTASUNA, EBOLUZIOA BIZKORKETARAKO GELAN, ALTA EMATERAKOAN, PROBA OSAGARRIAK, DIAGNOSTIKOA eta TRATAMENDUA ETA GOMENDIOAK.
- **AZPIATALAK (berdea):** BIHOTZAREN AUSKULTAZIOA, BIRIKEN AUSKULTAZIOA, SABELALDEA, GILTZURRUNAK, BIODIMIKIA, HEMOGRAMA eta KOAGULAZIOA.

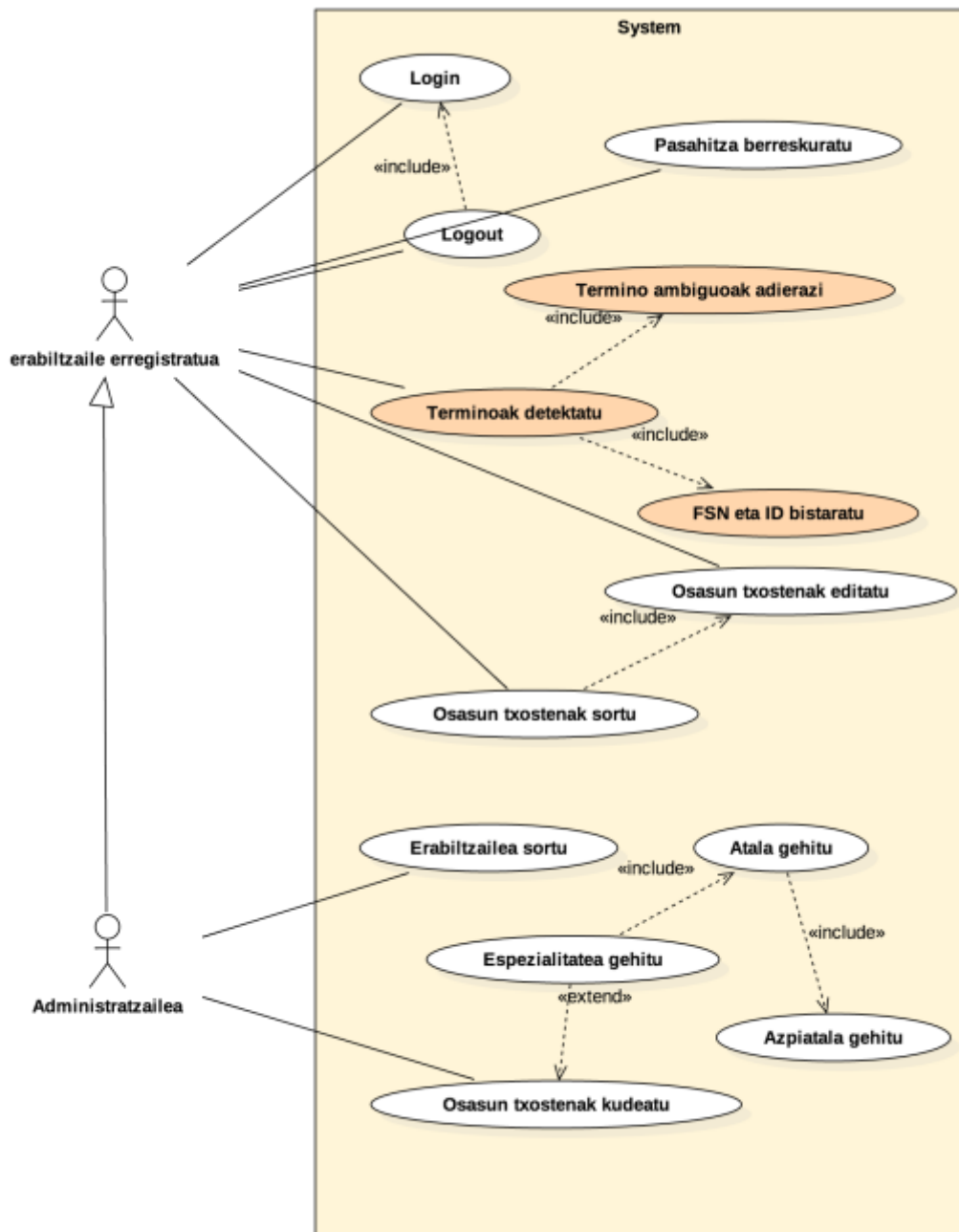
Behin bi sistemen klase-diagramak eginda eta *framework*-aren, aplikazioaren egituraz hitz egin ondoren, honen erabilpen-kasuei buruz hitz egingo dugu.

## 5.4 Erabilpen-kasuak

Azalpen bat da prozesu bat bete ahal izateko. Horretarako urrats edo ekintza batzuk egin beharko dira. Parte hartzen duten pertsonaiek aktoreak deitzen zaie. Gure kasuan aktore hauek administratzailea eta erabiltzaile erregistratua izango dira. Aktore bakoitzak erabilpen kasu jakin batzuk egin ahal izango ditu. Hau da, erabiltzaile batek (kasu honetan) ezin izango du espezialitate bat sortu ez duelako hau egiteko baimenik.

Bukatzeko, erabilpen-kasuak software ingeniartzaren inguruan interakzio sekuentzia batzuk dira sistema eta aktoreen artean garatuko direnak.

Hona hemen dauden erabilpen-kasuak (5.6 irudia). Laranja ikus ditzakegun erabilpen-kasuak proiektu honetan sortuak izan direnak dira eta txuriz aurretik sortuak zeudenak. Proiektu honentzako sortuak izan diren hiru erabilpen-kasuak hobeto azalduko ditugu eta aurretik sortuak zeudenak laburpen bat soilik emango da.



Irudia 5. 6: Erabilpen-kasuak.

## ERABILPEN-KASUEN AZALPEN LABURRAK:

**Erabiltzaile erregistratua:** zortzi erabilpen-kasu ditu

*Login egin:* erabiltzaileak aplikaziora sartzeko bere erabiltzaile izena eta pasahitza behar izango du. Behin bi hauek ondo sartu izan direla, erabiltzailea sistemaren barruan egongo da.

*Logout:* erabiltzailea irten *link*-aren gainean *klik* egiterakoan sistemaren kanpo egongo da.

*Pasahitza berreskuratu:* erabiltzailea (oraindik ez duena *login* egin) ez badu bere pasahitza gogoratzen pasahitza berreskuratzeko aukera edukiko du. Horretarako, bere *email*-a idatzi beharko du eta korreetik pasahitza berreskuratzeko egin beharrekoa izango du.

*Osasun-txostenak sortu:* Osasun-txostenak sortu ahal izango ditu behin sisteman *login* egin ondoren. Hemen osasun-txostenaren hainbat eremuetan idatzi behar izango du (atal eta azpiatal) espezialitatea aukeratu ondoren.

*Osasun-txostenak editatu:* Osasun-txosten bat edo batzuk sortu ondoren, hauek editatu ahal izango ditu eta sisteman moldaketa gordeko da.

### **1.-Terminoak detektatu:**

**LABURPENA:** behin *login* eginda eta atal edo azpiataletan idatzi ondoren, sistemak hainbat termino detektatuko ditu. Termino mota bakoitza kolore batean agertuko da hierarkia mota esleituz.

**AURREBALDINTZA:** saioa hasita izan.

**POSTBALDINTZA:** erabiltzaileak idatzi duen testutik terminoak detektatzen dira kolorezko etiketen bidez.

**GERTAERA-FLUXU NORMALA:**

- **Erabiltzailea:** espezialitate bat aukeratzen du.
- **Erabiltzailea:** beharrezkoak diren atazak eta azpiatazak betetzen ditu.
- **Sistema:** terminoak detektatzen ditu erabiltzaileak kurtsorea testu eremutik kentzerakoan.

**GERTAERA FLUXU ALTERNATIBOA:**

- **Erabiltzailea:** euskarak ez dituen berezko karaktereak (ñ, è, ë,...) erabiltzaileak erabili ditu.
  - o **Sistema:** *popUp* leiho bat bueltatuko du karaktere hauek ezin direla erabiliz esanez.

### **2.- Termino anbiguoak adierazi:**

**LABURPENA:** terminoak detektatu ondoren hainbat termino azpimarratuak agertuko dira. Termino anbigua dela esan nahi du.

**AURREBALDINTZA:** saioa hasita izan eta hainbat termino detektatu izana.

**POSTBALDINTZA:** anbiguoak diren terminoak azpimarratuak agertuko dira.

**GERTAERA FLUXU NORMALA:**

- **Sistema:** terminoaren analisia egingo du bilatuz ea SNOMED identifikatzaile bat baino gehiago dauden.

**GERTAERA FLUXU ALTERNATIBOA:**



Ez dago.

**3.-FSN eta ID bistaratu:**

**LABURPENA:** terminoak detektatu eta hauen gainean klikatu ondoren (anbiguoak edo ez) *popUp* leiho bat agertuko da, hauen SNOMED ID eta FSN-ak bueltatuz.

**AURREBALDINTZA:** saioa hasita izan eta hainbat termino detektatu izana.

**POSTBALDINTZA:** *PopUp* leiho berria agertzea terminoari dagokion SNOMED ID eta FSN-ekin

**GERTAERA FLUXU NORMALA:**

- **Erabiltzailea:** etiketatuta dagoen termino baten gainean klikatu.
- **Sistema:** termino horri dagozkion SNOMED ID eta FSNak bistaratu.

**GERTAERA FLUXU ALTERNATIBOA:**

Ez dago.

**Administratzailea:** hamahiru erabilpen-kasu ditu (berek bost eta erabiltzaile erregistratuaren zortziak).

*Erabiltzailea sortu:* administratzaileak erabiltzaileak sortu ahal izango ditu, nahi dituen pribilegioak emanaz. Administratzailea beste administratzaileak sortu ahal izango ditu ere.

*Osasun-txostenak kudeatu:* Administratzaileak osasun-txostenak kudeatu ahal zaingo ditu. Hau da, espezialitate berriak, espezialitate horien atalak eta atal hauei azpialak berriak gehitu ditzake.

*Espezialitatea gehitu:* espezialitate berriak gehitu ahal izango ditu (aparatu digestiboa, adibidez...).

*Atala gehitu:* espezialitate horiei atal berriak gehitu ahal izango dizkio (ospitaleratzeko/kontsultarako arrazoia, ...).

*Azpialak gehitu:* atalei azpialak gehitu ahal izango dizkio (bihotzaren auskultazioa edo biriken auskultazioa).

Behin bi sistemen eta aplikazioaren analisisa eginda, hiru hauek garatzeko egin den inplementazioari buruz hitz egingo dugu kode zatiak erakutsiz prozedura hobeto ulertzeko.

# 6

---

---

## Implementazioa

Atal honetan produktuaren implementazioa eta hartu diren erabakiak era justifikatu batean azalduko dira. Esan den bezala, med\_en sistema hobetuko dugu (6.1, 6.2 eta 6.3 azpiatalak) eta Med\_eu sistema sortuko dugu Med\_en sisteman oinarrituta (6.3, 6.4 eta 6.5 azpiatalak). Gainera, OsatEUS deituriko aplikazioa osatuko dugu (6.6 azpiatala) Med\_eu sistemako AnaMed\_eu analizatzailea erabiliz. Erabiltzaileak idazten duen testuaren terminoak detektatu eta etiketa kolorea esleituko zaio terminoei (bakunak, anitzak nahiz anbiguoak), bukaerako emaitza hau izanez. Bukatzeko, sortu diren arazoak eta hauen konponbideak azalduko ditugu (6.7 azpiatala).

## 6.1 Objektuen sorkuntza

Analisi atalean (5. Atala) esan den bezala, lau objektu berri sortu ditugu: *Concept* (kontzeptua), *Relationship* (erlazioa), *Descriptions* (deskribapenak) eta *Languaje* (lengoaia)

Hona hemen *Concept* klasea. Klase bakoitzean `__init__` deritzon funtzio pribatua gehitu da dagozkien atributuekin objektua hasieratu ahal izateko. Atributu bakoitza `self.` aurrizkiarekin egongo da adierazita (6.1 testu-kutxa).

```
class Kontzeptua:

    def __init__(self, cId, isP, fsn = "", snoId = "",
synDesc = "", preDesc = ""):
        self.conceptId = cId
        self.isprimitive = isP
        self.fullySpecifiedName = fsn
        self.snomedId = snoId
        self.synonymDesc= synDesc
        self.preferredDesc = preDesc
```

*Testu-kutxa 6. 1: kontzeptu objektuaren funtzio hasieratzailea.*

Lau objektuek dituzten atributu guztiak `get` eta `set` funtzioen bidez hasieratu behar dira (6.2 testu-kutxan ez dira atributu guztiak idatzi azpi atal hau luzeegia ez izateko).

```
def setConceptID(self, cId):
    self.conceptId = cId

def getConceptID(self):
    return self.conceptId

def setIsPrimitive(self, isprimitive):
    self.isprimitive = isP

def getIsPrimitive(self):
    return self.isprimitive

...
```

*Testu-kutxa 6. 2: kontzeptu objektuaren get eta set funtzioak.*

Beste hiru objektuekin berdina egin dugu, bakoitzari dagozkion atributuak esleituz (ikusi 6.1 eta 6.1 testu-kutxak).

Behin objektuak sortuak izan direla, aurretik sortuak zeuden funtzioak (*LanguageList*, *ConceptList*, *DescriptionList* eta *RelationshipList*-eko `__init__` funtzioak) moldatuak izan dira.

Sortu berri diren objektuak deitu behar izan ditugu hauek hasieratu ahal izateko eta dagozkien parametroak pasa. Hau egin da `languageJaso`, `erlazioaJaso`, `kontzeptuaJaso` eta `deskribapenaJaso` funtzioetan. Hona hemen `kontzeptuaJaso`-ko erredua:

```
def kontzeptuaJaso(self, line):
    eremuak = line.strip().split('\t')
    ...
    concept = Kontzeptua(eremuak[0], eremuak[5],
eremuak[2], eremuak[4]) #objetuari balioak pasa
```

*Testu-kutxa 6. 3: Hasieratze funtzioari deitu.*

Set-etan bihurtzeko:

```
c.fullySpecifiedName = d.term -> c.setFSN(d.term)
```

*Testu-kutxa 6. 4: set funtzioaren bihurketa.*

Get-etan bihurtzeko:

```
d.conceptId -> d.getConceptId()
```

*Testu-kutxa 6. 5: get funtzioaren bihurketa.*

Hau eginda, `klaseak.py` eta `klaseak_eu.py` hiztegiatik objektuetara moldatuta izango genuke finkatutako helburu bat betez. Beraz, kontzeptu, erlazio, deskribapen eta hizkuntzak beste modu batera biltegiatzea lortu dugu. Behin hau eginda (graduan objektuekin lan egiten ikasi dugulako), `Med_eu` sistema hobetzen hasiko gara bilaketak lemaren arabera eginez.

## 6.2 Terminoak identifikatzen lemaren arabera

Ingeleseko bertsioan ez bezala, euskarazko bertsioan terminoen sctId-a (SNOMED identifikatzailea) modu egoki batean aurkitu ahal izateko lema<sup>5</sup> behar izango ditugu.

Ingeleseko bertsioan ez du berezko garrantzia deskribapenak (edo terminoak) eta hauen lema askotan berdinak izango direlako (ikusi 6.1 taula).

Terminoak	Lema
Down syndrome	Down syndrome
Allergy	allergy
Sore throat	sore throat
Head ache	head ache

Taula 6. 1: terminoak eta lemaren arteko ezberdintasuna ingelesez

Ingelesaren kasuan ezaugarri hau dela eta, bilaketak deskribapenaren edo lemaren arabera egin daitezke, hauen artean ez baitago diferentzia handirik (ikusi 6.1 taula). Euskarazko bertsioan, aldiz, honen arabera bilaketak egitea berebiziko garrantzia izango du euskarak duen morfologiarengatik. Hau da, etxetik, etxera edo etxeak hitz desberdinak izan arren, lema berdina edukiko dute etxe hain zuzen ere (6.2 taulan adibide gehiago ikus ditzakegu).

Terminoak	Lema
Downen sindromea	Down sindrome
Down-en sindromea	
Down sindromea	
Downen sindromearekin	
Alergiarik	Alergia
Alergia	
Alergiarekin	
Buruko mina	Buru min
Buruko minarekin	

Taula 6. 2: terminoak eta lemaren arteko ezberdintasuna euskaraz.

<sup>5</sup> Unitate autonomoa hizkuntza baten lexikoari dagokiona

Beraz, lemaren arabera bilaketak egin behar izango ditugu bilaketak arrakastatsuak izateko. Horretarako, hurrengo aldaketa hauek egin dira deskribapen (termino) bakoitzaren lemari SNOMED Identifikatzailea ezartzeko. SNOMED Identifikatzailea lortu da 6.1 azpiataleko `get (getConceptId())` funtzio baten bidez. Aldaketa hauek `DescriptionList()` klaseko `__init__` funtzioan egin dira:

```

self.lemArabera =[] #aldagaia hasieratzen
...
#for-aren bidez termino bakoitzaren lemma guztiak
lortuko ditugu eta kateatuko ditugu.
lemmak = ""
sentence = result['sentences']
for sentence in result["sentences"]:
    lemma = sentence[1]['Lemma']
    lemmak +=str(lemma)+" "
#azkeneko karakterea ezabatu (" ")
lemmak = lemmak[:-1]
#kontzeptuaren (terminoa) ida lortu
cId = d.getConceptId()
#lemeei kontzeptuen ID-aren arabera identifikatu
self.lemArabera[lemmak] = cId

```

*Testu-kutxa 6. 6: terminoen lemei dagozkien SNOMED IDak esleitu.*

`self.lemArabera` aldagaiak termino jakin baten lema eta termino horri dagokion SNOMED Identifikatzailea (`'Lema': 'SNOMED ID'`) bueltatuko luke:

```

{'kordegabetu': '271594007', 'perikolangitis': '111373008',
'linfoma histiozitiko egia': '109988003', 'ondoez jarri':
'271594007', 'lipofuszinosi zeroide neuronal': '61663001',
'huts hurrengo': '271594007', 'makal': '271594007',
'dermatitis esfoliatiboa': '396350005', 'txiki':
'271594007', 'pattal': '271594007', 'larruazal muzinosia':
'238946004', 'alditxar egin': '271594007',
'hiperaldosteronismo primario': '237742008', 'indargabe':
'271594007', 'lo ( ara ) trastorno , zehaztugabe':
'230498005', 'konorte-galera': '271594007'}

```

*Testu-kutxa 6. 7: lema bakoitzak edukiko duten SNOMED IDak.*

Ikus dezakegu nola lema bakoitzak SNOMED Identifikatzaile bat izango duen gutxienez. Behin hau eginda, erabiltzaileak idazten dituen terminoen lemak lortu behar izango ditugu (termino luzeena lortuz). Horretarako, hurrengo hau gehitu zaio (6.8 testu-kutxako) bezeroaren `snomedIdentifikatu()` funtzioan (beltzez agertzen dira egin diren aldaketak eta gehikuntzak).

```

for token in tagged:
    forma,info = token
    formak.append(forma)
    llag = len(fArray)
    k += 1
    fArray.append(forma)
    #erabiltzailearen idatzi duen terminoaren lemmak
    kateatuko ditugu array batean
    lArray.append(info["Lemma"])
    for i in range(0,llag):
        #l aldagaian lArrayko elementua kateatu string
        batean
            l = lArray[i] + " " + info["Lemma"]
            f = fArray[i] + " " + forma
            lArray[i] = l
            fArray[i] = f
            kodT = des.desc2sct('', l.lower())

```

Testu-kutxa 6. 8: termino anitzetatik SNOMED ID lortu.

Hau da, lemmak kateatuz (append() funtzioa) termino luzeenaren SNOMED Identifikatzailea lortuko gendake (existitzen bada) desc2sct() funtzioaren bitartez. desc2sct() moldatu beharko dugu lemArabera listan SNOMED ID-ak lortzeko. Horretarako, l.lower() parametrotzat pasako zaio (beltzez agertzen dira egin diren aldaketak).

```

def desc2sct(self,desc,lemma):
    return json.dumps(self.desZer.kodeaLortu (desc,lemma))

#get funtzioaren bitartez lemma sartzerakoan honen Ida
buletatuko du
def kodeaLortu(self,forma,lemma):
    if lemma:
        return self.lemArabera.get(lemma, '')
    else:
        return self.lemArabera.get(forma, '')

```

Testu-kutxa 6. 9: lemmen arabera SNOMED ID-a lortzeko funtzioa.

Hona hemen prozesuaren adibide bat:

1.- erabiltzaileak sartutako terminoa: *hiperaldosteronismo primarioa*

2.- bi hitzen lemmak lortu: *hiperaldosterismo, primario*

3.- lemmak konkatatu: *hiperaldosterismo primario*

4.- begiratu lema hau *lemArabera* aldagaian aurkitzen den (ikusi 6.7 testu-kutxa).

5.- *lemArabera* aldagaian dagoenez, terminoa da eta 237742008 SNOMED identifikatzailea lortuko luke (6.9 testu-kutxaren funtzioen ondorioz).

Hau guztia egin ondoren, terminoak lemaren arabera bilatzea lortuko genuke eta euskararako bertsiorako oso lagungarria izango da bilaketak egiterakoan. Deskribapen baten lema lortzeko hitzaren analisia egin behar dugunez (*eustagger-lite* eta *Stanford* analizatzaileei hainbat deiak egin behar dizkiegu) sistema asko moteltzen da. Hau dela eta, datu-base bat sortu behar dugu arazo honi aurre egiteko. Beraz, hurrengo atalean hainbat datu-base moten analisia egingo da aproposena aukeratuz.

## 6.3 Datu-basea

Ingeleseko bertsioan deskribapenak kargatzerakoan 500.000 deskribapen baino gehiago zeudenez, denbora asko behar zuen.

Segunduro bost edo lau deskribapen kargatzen ziren. Hasiera batean, 512.724 deskribapen genituen eta demagun kasurik onenean bost deskribapen segunduro kargatzen direla:

$$\frac{512724}{5} = 102544,8 \text{ segundu} \approx 29 \text{ ordu}$$

*Testu-kutxa 6. 10: ingeleseko informazio lexikoa duen fitxategia sortzeko behar izan den denbora.*

29 ordu pasako lituzke (ikusi 6.10 testu-kutxa) *MedTerm\_en* sistemak deskribapenak soilik kargatzen, eta hau ez da bideragarria. Zailtasun hau ekiditeko, datu-base bat sortu dugu (lehendik ez zegoena, deskribapenak *RF1* fitxategi batetik datuak atzitzen zirelako *.txt* formatua duena), eta horretarako hainbat baliabideen analisia egin dugu zein den egokiena aukeratzeko, hots, zein sistema den azkarren.

Euskarako bertsioan aldiz, deskribapen kopuru txikiagoa izango da (222504) baino hauen analisia motelagoa da *eustagger\_lite* analizatzailea motelagoa baita. Sei segunduro bost deskribapen idazten dira, hau da, 1.2 segundo behar dira deskribapen bakarra idazteko.



Eustagger\_lite analizatzailea erabili behar dugu lematizatzailea delako (eta bilaketak lemen arabera egingo direlako).

Demagun 1.2 segundo behar dela analisi bat egiteko, beraz:

$$222504 \times 1.2 = 267004,8 \text{ segundo} \approx 74 \text{ ordu}$$

*Testu-kutxa 6. 11: : euskarazko informazio lexikoa duen fitxategia sortzeko behar izan den denbora.*

Beraz, bi datu-base sortuko dira ingeleseko eta euskarazko MedTerm zerbitzariak deskribapenak azkarrago kargatzeko eta honekin batera, bi zerbitzarien martxan jartzea azkarragoa izateko.

Lehendik ataza hau azkartzeko erabili daitezkeen errekurtso posibleak adierazi (kusi 5.3 taula) eta zer diren deskribatu dugu (hasieran sarrera motz batekin eta ondoren baliabideen deskribapen sakonagoa eginez). Hau egin ondoren, denen arteko konparaketak (ikus 5.4 taula) egin dira:

**ERREKURTZO POSIBLEAK:** TSV, MySQL Datu-basea, *Pickle* liburutegia eta JSON.

BALIABIDEA	ZER DA?	ABANTAILAK	DESABANTAILAK
TSV	TXT fitxategi tabulatua.	- Sinplea. - Irakurgarritasuna. - Excel-en erabilgarri	-
MySQL DATU BASEA	Zerikusia duten datuen bilduma, ondoren eskuratu ahal izateko.	- Informazioa egituratzeko aproposa.	- Inplementazio konplexua. - Mantenimendu jarraitua.
PICKLE LIBURUTEGIA	Python-entzako objektu serializatzailea.	- Python-entzako aproposa. - Erabilera sinplea. - Beste programetan erabilgarria.	- Motela - Ez ikusgarria (kode bitarra)
JSON	Testu hutsezko formatua datu egitura sinpleak eta array asoziatiboak	- Atzitzeko azkarra. - Lengoaiei independentea.	-

errepresentatzeko  
erabiltzen da.

Taula 6. 3: baliabideen sarrera txikia.

Baliabidetan sakonduz:

**TSV (Support.google.com., 2016):**

Fitxategi hauek (ingelesez, *Tab Separated Values*) dokumentu mota batzuk dira non datuak errepresentatzen diren taulen egiturekin. Zutabeak tabulazioen (<tab>) bidez banatzen dira eta lerroak, aldiz, lerro jauzien bidez (\n). Beraz, TSV formatua hurrengo erregelak jarraitzen ditu:

- Sarrera bakoitzak lerro bakarra dauka.
- Lehengo lerroa goiburua da non eremu bakoitzaren balioa definitzen den.
- Eremuak tabulazioekin banatzen dira.
- Eremuek datuak eduki behar dute (zifra edo testua).
- Lerro guztiek eremu kopuru berdina eduki behar duzte.

Hau da, hurrengo hau idatziko balitz:

```
Izenburua <tab>Egilea<tab>baldintza  
Pride and Prejudice<tab>JaneAusten<tab>used  
The Art of War<tab>Sun Tzu<tab>used
```

Testu-kutxa 6. 12: TSV fitxategia sortzerakoan duen itxura.

Hau ikusiko zen testu editore batean:

<b>Title</b>	<b>Author</b>	<b>Baldintza</b>
Pride and Prejudice	Jane Austen	used
The Art of War	Sun Tzu	used

Testu-kutxa 6. 13: TSV fitxategia sortu ondorengo itxura.

Beraz, ikusten denez ulertzeko oso erraza izango zen eta begi bistara oso garbia da.

**MySQL DATU-BASEA:**

Hainbat datu base mota egon arren, MySQL datu-basea baliabide posibletzat erabaki da aukeratzea graduan gehien joratu izan den datu-basea

izan delako eta Python-entzako baliagarri dagoelako. Horretarako, programetan 5.14 testu kutzako kodea idatzi beharko da instalatzeko.

```
import MySQLdb.
```

*Testu-kutxa 6. 14: Pythonen MySQL erabiltzeko inportatu behar den paketea.*

Behin hau eginda, konexioa ezarri eta jadanik kontsultak egin ahal izango dira `execute()` aginduaren bitartez (6.15 testu-kutzaren hirugarren agindua `cursor.execute("SELECT VERSION()")`).

```
#!/usr/bin/python
import MySQLdb

# Datu-base konexioa ireki
db =
MySQLdb.connect("localhost","testuser","test123","TESTDB" )
# Kurtsore objetua prestatu cursor() metodoa erabiliz.
cursor = db.cursor()
# SQL querya exekutatu.
cursor.execute("SELECT VERSION()")
# Zerbitzariatik deskonektatu
db.close()
```

*Testu-kutxa 6. 15: MySQL datu basearekin konexioa egiteko eman behar diren urratsak.*

Ondoren, taulak sortu, tuplak gehitu etab. egin beharko ziren (kasu honetan deskribapenezako tupla asko egongo ziren).

### **PICKLE LIBURUTEGIA (Pyspanishdoc.sourceforge.net., 2016):**

Algoritmo erraza inplementatu eta objektuak serializatzen ditu. Hau da, objektuak *byte* kate batean bihurtzen ditu (eta alderantziz). Beraz, objektu konplexuak *byte*-tara bihurtzeko (*pickleatzea* deitzen zaiona) `dump()` aginduaren bitartez (testu fitxategian gehitu ditzakegu objektuak).

`x` objektua eta `f` fitxategia edukita `dump()` egingo dugu *pickleatze*ko:

```
pickle.dump(x, f)
```

*Testu-kutxa 6. 16.: x objektua f fitxategian gehitu.*

Kodea *despickleatze*ko, aldiz, hurrengo agindua erabiliko dugu:

```
x = pickle.load(f)
```

*Testu-kutxa 6. 17: f fitxategia x objektuan kargatu.*

Beraz, ikusten den moduan erabiltzeko erraza da.

### **JSON:**

- Datuen elkartruckerako formatu arina da (Json.org, 2016) Testu hutsezko formatu hau datu egitura sinpleak eta *array* asoziatiboak errepresentatzeko erabiltzen da. JSON fitxategiak *.json*-en bidez adieraziko ditugu.

JSON-ak bitan egituratuta daude:

- Hiztegieta non izen/balio bikoteko zerrendak dauden. Hiztegiak giltzen bidez (“{” eta ”}”) identifikatzen dira eta kideez osatuta egongo dira. Kide bakoitzak izen errepikaezin bat edukiko du eta “:” izango du jarraian, ondoren balio batekin izendatzeko. Bikoteak “,” bidez bananduko dira.
- *Array*-etan non balio ordenatuen listak diren. Hauek kortxeteen bidez (“[” eta ”]”) identifikatzen dira eta hauen barruan hainbat elementu egongo dira.

JSON-en egitura:

```
hiztegia
  {}
  { kideak }
kideak
  bikote
  bikote, kideak
bikote
  string : balio
array
  []
  [ elementuak ]
elementuak
  balio
  balio, elementuak
balio
  string = "" edo "karakterreak"
  number = -2, -1.333, 0, 1, 2, 3.5, 3.99, ...
  objektua
  array
  true
  false
  null
```

Testu-kutxa 6. 18: JSON-en egitura.

Pickle liburutegiarekin bezala, `load()` (`.json` fitxategian dagoena Python objektu batean bihurtuko du) eta `dump()` (Python objektua JSON objektu batean bihurtu, `.json` fitxategian gordetzeko alegia) funtzioen bidez baliatuko gara JSON-ak sortzeko.

### 6.3.1 Baliabidearen aukeraketa

Behin errekurtsiak azalduta, taula bat sortuko dugu (6.4 taula) produktura hobeto moldatzen den datu basea aukeratuz:

ERREKURTSIAK	IRABAZLEA	ARRAZOIA
DB vs Pickle	DB	<ul style="list-style-type: none"> <li>- <i>Pickle</i>-k berezko formatua duenez, ez da erabilgarria beste tresnerako, beraz, irakurketak eta idazketak soilik egin ahal izango dira (ezingo da ikusi).</li> <li>- DB-ak formatu estandarra dutenez datuak manipulatzeko errazago izango da.</li> </ul>
DB vs JSON	JSON	<ul style="list-style-type: none"> <li>- Kasu honetan, JSON egokiagoa izango da, erabiltzen ditugun datuak jadanik testua JSON moduan ditugulako eta eremuak ez direlako eraldatuak izango.</li> <li>- Objektuen eremuak maiz aldatu, kendu edo gehitu behar izango balira, MySQL egokiena izango litzateke (Programmers.stackexchange.com, 2016).</li> </ul>
TSV vs JSON	JSON	<ul style="list-style-type: none"> <li>- Deskribapenak JSON fitxategi batera moldatu daitezkeenez JSON erabiltzea erabaki dugu TSV-a ordez.</li> </ul>
Pickle vs JSON	JSON	<ul style="list-style-type: none"> <li>- <i>Pickle</i> liburutegia baztertuta geratzen da aurreko arrazoiarengatik.</li> <li>- JSONek formatu estandarra dutenez (DBak bezala) datuak manipulatzeko errazago izango da.</li> <li>- JSON 25 alditan azkarrago da irakurketak egiterakoan (<code>load</code>) eta 15-etan idazterakoan (<code>dump</code>) (Kovshenin, 2010).</li> </ul>
DB vs TSV	DB	<ul style="list-style-type: none"> <li>- MySQL-en bidez TSV edo CSV fitxategiak sortu daitezke MySQL taulan bidez.</li> </ul>

Taula 6. 4: balibideen aukeraketa.

Beraz, 6.4 taulan ikusten den bezala, produktuarentzako JSON datu-basea aukerarik hoberena izango litzateke.

### 6.3.1.1 JavaScript Object Notation

Baliabide hau erabilia nahiko aldaketa gutxi egin behar izan dira Stanford zerbitzarian. `load()` eta `dump()` funtzioen bidez baliatuko gara `.json` fitxategia betetzeko datuak kargatzen diren heinean. Stanford analizatzaileak JSON formatuan bueltatzen ditu terminoen analisiak eta honengatik JSON datu-basea aukeratu da (lana erraztuz). Python-ek eskaintzen duen `open()` funtzioaren bitartez fitxategian idatzi ahal izango dugu JSON formatuan.

Beraz, hurrengo hau egingo dugu: `TermMed` zerbitzarian lehenengo aldiz martxan jartzerakoan informazio lexikoa duen fitxategia sortuko du, fitxategi hau hutsa dagoelako (zerbitzari 29 orduz martxan eduki behar da hau betetzeko, ikusi 6.10 testu-kutxa). Hurrengo aldietarako hau ez dugu berriro egin behar fitxategi lexikoa deskribapenez beteta dagoelako eta beraz, `TermMed` zerbitzariak martxan segituan jarriko dira. Fitxategietan deskribapen guztien analisi linguistiko osoa agertuko da.

#### JSON fitxategia sortzen:

Baliabideen direktorioan hainbat fitxategi izango ditugu. Bost fitxategi sortuko ditugu (non bi fitxategi `analisiak_en.json` eta `analisiak_eu.json` kopiak izango diren):

- `proba.json` -> 10000 deskribapen izango ditu sorkuntzarekin probak egiteko. JSON hau betetzeko 36 minutu behar izango ditu gutxi gora behera eta ingeleseko bertsiorako sortuta izango da. Standord analizatzailea erabili da hau sortzeko.
- `analisiak_en.json` -> ingelesez dauden deskribapen guztiak hemen azalduko dira.
- `analisiak_eu.json` -> euskaraz dauden deskribapen guztiak hemen azalduko dira.
- `analisiak_en_kopia.json` -> `analisiak_en.json`-aren kopia.

- `analisiak_eu_kopia.json` -> `analisiak_eu.json`-aren kopia.

Lehenengo hiru fitxategi hauek (beste biak kopia hutsak dira) “{}” karakterea izango dute beraien barnean, hasiera batean deskribapenak lista batean sartuko direla adierazteko (ikusi 6.23 testu-kutxa). Ikusi 6.19 testu-kutxa.

Hau egin ondoren `corenlp.py` eta `klaseak.py` fitxategietan aldaketak egingo dira (euskarako bertsioan aldaketa berdinak egin behar dira `wrapper_eustagger.py` eta `klaseak_eu.py` fitxategietan):

Informazio lexikoa izango duen fitxategiaren *path*-a idatziko dugu `ANALYSIS_DIR` aldagai globalean:

```
ANALYSIS_DIR =
os.environ.get("ANALYSIS", "/sc01a4/users/aelorz003/KBP/bali
abideak")
```

*Testu-kutxa 6. 19: ANALYSIS\_DIR aldagai globalaren sorketa.*

Behin hau eginda, bi aldagai berri sortuko ditugu non lehengoan informazio lexikoa fitxategiaren *path* guztia agertuko den, adierazteko non dagoen erabiliko dugun fitxategia (kasu honetan `analisiak_en.json` izango da). Bigarren aldagaian, aldiz, jadanik sortuak izan diren deskribapen guztiak sartuko ditugu. Honen izena `self.pre_loaded_analysis_dict` izango da (6.20 testu-kutxa).

```
def __init__(self, corenlp_path=DIRECTORY, memory="3g",
properties='default.properties', serving = False,
analysis_path=ANALYSIS_DIR) :

    #json fitxategia kargatu
    self.pre_analysis = analysis_path +
"/analisiak_en.json"
    #analisiak_eu.json dagoen berreskuratu
    with open(self.pre_analysis, encoding = 'utf-8') as f:
        self.pre_loaded_analysis_dict = json.load(f)
```

*Testu-kutxa 6. 20: JSON fitxategiaren kargateka self.pre\_loaded\_analysis\_dict aldagaiaren (beltzez, aurretik aipatutako aldagaiak)*

`parse()` funtzioa 512724 alditan deitua izango da (deskribapen kopurua) eta `self.pre_loded_analysis_dict` hiztegiaren `results` balioa (*stanford*-ek egindako analisia) eta `to_send` gakoa (deskribapenaren izena) gordeko dira hurrenag formatuarekin `{ 'to_send' : 'results' }`.

`open()` ('w' letra gehituz, fitxategian idatzi ahal izango dugu) funtzioa eta `self.pre_analysis`-ren bitartez `self.pre_loaded_analysis_dict` dauden deskribapen guztiak fitxategian idatziko ditugu.

Fitxategian deskribapenak 20.000-naka idatziko dira segurtasun maila bat ezarriz (dena batera ez egiteko, argia joaten bada lan guztia ez galtzeko, ...). Ikusi 6.21 testu-kutxa:

```
def _parse(self, text):
    ...
    self.pre_loaded_analysis_dict[to_send] = results

    #20000naka pre_loaded_analysis_dict-en daogena idatzi
    fitxategian
    if len(self.pre_loaded_analysis_dict) % 20000 == 0 or
    len(self.pre_loaded_analysis_dict) == 498613:
        self.storeAnalysis()
    return results

#JSON fitxategian idatzi
def storeAnalysis(self):
    ...
    with open(self.pre_analysis, "w", encoding = 'utf-
8') as f:
        json.dump(self.pre_loaded_analysis_dict, f)
```

Testu-kutxa 6. 21: 20.000 deskribapenero fitxategian idatzi.

Aurretik esan den bezala, `dump()` funtzioaren bitartez `self.pre_loaded_analysis_dict` dagoena `f` aldagaian idatziko dugu JSON objektua bihurtuz (`results` aldagaian *string* moduan aurkitzen baitzen).

`klaseak.py` fitxategiaren bitartez, lortuko da jakitea ea terminoa JSON-ean dagoen edo ez dagoen ikustea. Terminoa oraindik ez bada terminoen informazio lexikoa duen fitxategian aurkitzen (`if` baldintzaren arabera) *Stanford*-eko zerbitzariak (`corenlp.py` fitxategia) terminoa *parse*atuko du eta fitxategian gehituko da (6.22 testu-kutxan beltzez agertzen den kode zatia). Hau eginez, deskribapenen karga segituan egingo da (bi minutuko inguruko denboran).



```

self.pre_analysis =
"/sc01a4/users/aelorz003/KBP/baliabideak/analisiak_en_proba
.json"
with open(self.pre_analysis, encoding = 'utf-8') as f:
    self.pre_loaded_analisys_dict = json.load(f)
...

#begiratu ea terminoa jadanik JSON fitxategian dagoen
if d.term in self.pre_loaded_analisys_dict:
    result =
self.pre_loaded_analisys_dict[d.term]
else:
    nlp = StanfordNLP()
    result = nlp.parse(d.term)

```

Testu-kutxa 6. 22: konprobatu ea d.term terminoa fitxategian dagoen.

Ikusi 6.23 testu-kutxako terminoen informazio lexikoa duen .json fitxategia hartuko duen itxura (beltzez terminoa eta etzanez terminoaren informazio linguistikoa).

```

{"Lack of ossification of scapula": {"sentences": [{"text": "Lack
of ossification of scapula", "words": [{"Lack", {"PartOfSpeech":
"NN", "NamedEntityType": "O", "CharacterOffsetBegin": "0", "Lemma":
"lack", "CharacterOffsetEnd": "4"}], [{"of", {"PartOfSpeech": "IN",
"NamedEntityType": "O", "CharacterOffsetBegin": "5", "Lemma": "of",
"CharacterOffsetEnd": "7"}], [{"ossification", {"PartOfSpeech":
"NN", "NamedEntityType": "O", "CharacterOffsetBegin": "8", "Lemma":
"ossification", "CharacterOffsetEnd": "20"}], [{"of",
{"PartOfSpeech": "IN", "NamedEntityType": "O",
"CharacterOffsetBegin": "21", "Lemma": "of", "CharacterOffsetEnd":
"23"}], [{"scapula", {"PartOfSpeech": "NN", "NamedEntityType": "O",
"CharacterOffsetBegin": "24", "Lemma": "scapula",
"CharacterOffsetEnd": "31"}]]], "dependencies": [], "parsetree": [],
"indexeddependencies": []}}, {"32 French gauge": {"sentences":
[{"text": "32 French gauge", "words": [{"32", {"NamedEntityType":
"NUMBER", "NormalizedNamedEntityType": "32.0", "CharacterOffsetEnd":
"2", "PartOfSpeech": "CD", "CharacterOffsetBegin": "0", "Lemma":
"32"}], [{"French", {"PartOfSpeech": "JJ", "NamedEntityType": "MISC",
"CharacterOffsetBegin": "3", "Lemma": "french",
"CharacterOffsetEnd": "9"}], [{"gauge", {"PartOfSpeech": "NN",
"NamedEntityType": "O", "CharacterOffsetBegin": "10", "Lemma":
"gauge", "CharacterOffsetEnd": "15"}]]], "dependencies": [],
"parsetree": [], "indexeddependencies": []}}

```

Testu-kutxa 6. 23: informazio lexikoa duen fitxategiaren zati bat.

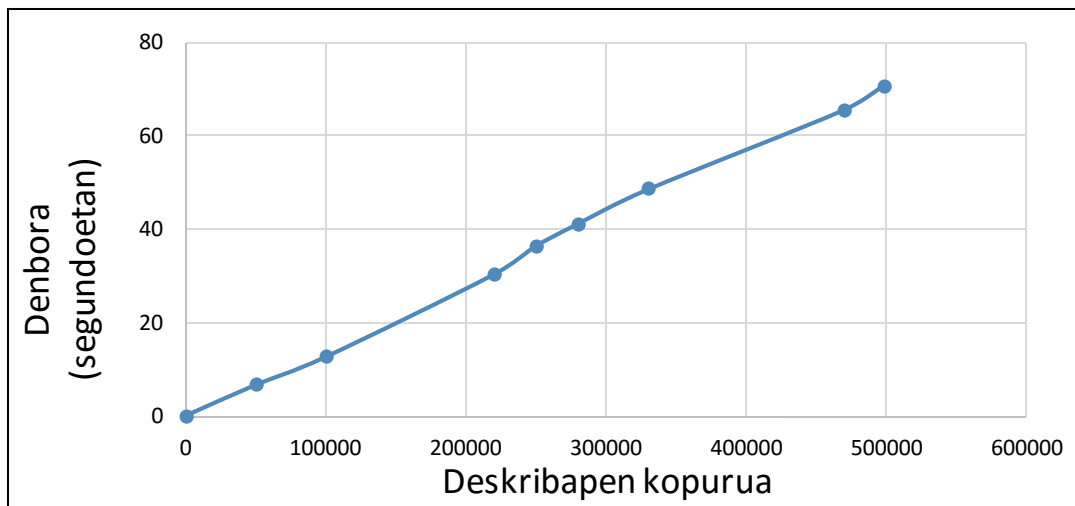
## dump()

Ikusi 6.5 eta 6.6 tauletan `dump()` aginduak fitxategian idazketak egiteko behar duen denbora (segundoetan). Denbora kopurua lortzeko `time()` paketea inportatu behar da beltzez agertzen den 6.24 testu-kutxako kodea idatziz).

```
def storeAnalysis(self):  
  
    #lortu hasierako ordua (momentu horretakoa)  
    start_time = time.time()  
    #fitxategian idatzi  
    with open(self.pre_analysis,"w", encoding = 'utf-8') as f:  
        json.dump(self.pre_loaded_analysis_dict,f)  
  
    #kenketa egin momentuko orduarekin eta aurretik lortu den  
    ordua (start_time)  
    print("Idazketa egiteko denbora:", time.time() - start_time)
```

Testu-kutxa 6. 24: idazketak egiteko behar den denbora lortu.

Bukaeran lortu izan den grafikoa MedTerm\_en sisteman (6.1 grafikoa).



Grafiko 6. 1: `dump()` aginduak `analisiak_en.json` fitxategian idazteko behar duen denbora.

Beste era batera erakusteko, ikusi 6.5 taula.

Desk. kop	Denbora (seg)	Desk. kop.	Denbora (seg)	Desk. kop	Denbora (seg)
0	0	220000	30,19	330000	48,38
50000	6,6	250000	36,23	470000	65,38

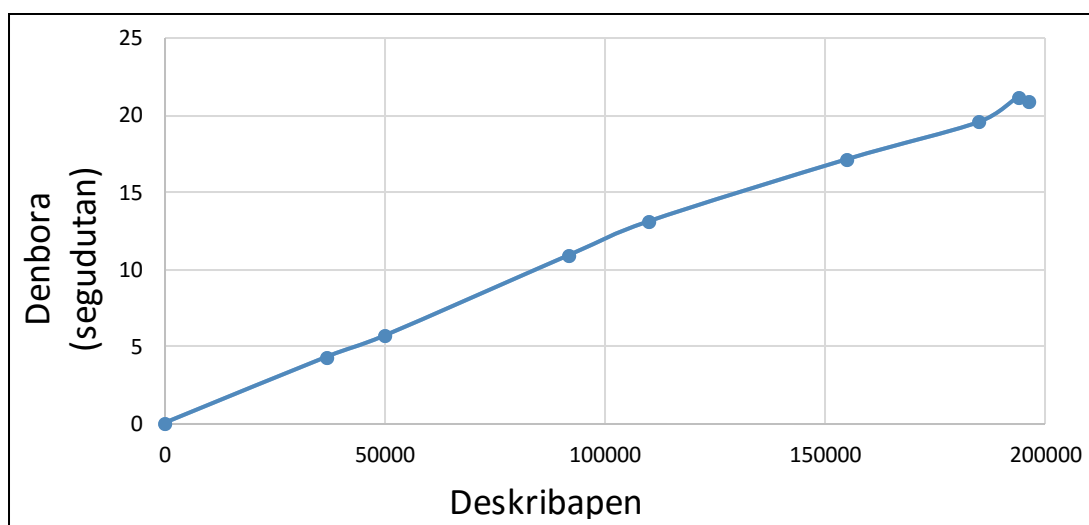
100000	12.58	280000	41,01	498613	70,61
--------	-------	--------	-------	--------	-------

Taula 6. 5: dump() aginduak analisiak\_en.json fitxategian idazteko behar duen denbora.

Beraz, dump() gutxi egin behar dira JSON-aren sorkuntza denbora gehiago ez inbertitzeko (29 ordu behar dira JSON-aren sorkuntza egiteko). 20.000 deskribapenera (gutxi gorabehera 72 minuturo) dump() egin da. Azkenean ez dira fitxategian honetan espero ziren termino guztiak gehituak izan (512.724), baizik eta 498.613. Gehituak izan ez direnak RF1 fitxategian errepikatuak zeudenak izan dira.

MedTerm\_eu sistemako analisiak\_eu.json fitxategian 196.229 deskribapen gehitu dira, RF1 fitxategian (.txt formatukoa) deskribapen asko errepikatuak zeudelako eta hauek ez dira gehituak izan. Hona hemen idazketak egiteko behar izan den denbora. Fitxategi hau sortu ondoren, deskribapen guztiak kargatzeko 57 segundo inguru behar izan dira:

dump() aginduak idazketak egiteko behar duen denbora euskarako bertsioan 6.6 taulan eta 6.1 grafikoan aurkitzen dira.



Grafiko 6. 2: dump() aginduak analisiak\_eu.json fitxategian idazteko behar duen denbora.

Desk. kop	Denbora (seg)	Desk. kop	Denbora (seg)	Desk. kop	Denbora (seg)
0	0	92000	10.91	185000	19.54
37000	4.31	110000	13.09	194000	21.16
50000	5.67	155000	17.12	196229	20.82

Taula 6. 6: dump() aginduak analisiak\_eu.json fitxategian idazteko behar duen denbora.

Beraz, fitxategi hauen sorkuntzaren ondorioz bi sistemak azkarrago kargatuko dira eta errepikatuak zeuden deskribapenak deuseztatuak izango dira.

### TERMINO ANBIGUOAK:

Euskaraz eta ingeleseko informazio lexikoa gordetzen diren fitxategien termino kopurua txikiagoa da bi RF1 fitxategiekin konparatuz. Honen azalpena erraza da: RF1 fitxategietan hainbat termino errepikatzen dira, eta beraz, hauen informazio lexikoa ez da berriro gordetzen. Hainbat alditan errepikatuak egon arren termino hauek dituzten SNOMED ID-ak (`sctId`) desberdinak dira, beraz, termino hauek anbiguoak direla (esanahi bat baino gehiago) esaten da. Hau da, “#” karakterea ikusten bada terminoaren analisia egiterakoan, termino honek anbiguo delako esanahi du. Adibidez:

```
Alergia-en SNOMED IDak: 609328004#408439002, beraz, RF1 fitxategian alergia terminoa bitan agertuko da.
```

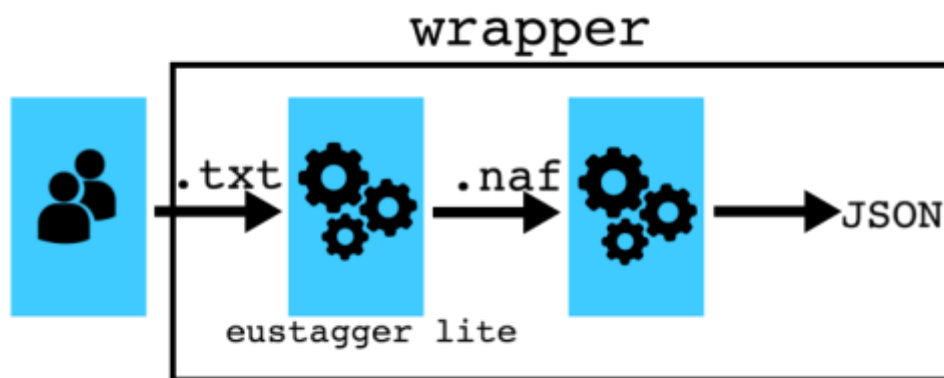
*Testu-kutxa 6. 25: alergia termino anbigua.*

Ingeleseko bertsioan hiru moldaketa hauek (objektuen sorketa, terminoen bilaketak egitea lehen araberako egitea eta datu-basearen sorketa, hau da, 6.1, 6.2 eta 6.3 azpiatalak) egin ondoren euskarazko bertsioa implementatzen hasiko gara. Horretarako, `eustagger-lite` erremintaz (4.3.2 azpiatala) baliatuko gara eta bueltatzen duen emaitzaren itxura aldatu beharko dugu `wrapper`-aren (ikusi 6.1 irudia) sorkuntzarekin.

## 6.4 Wrapper-ra

Aurretik esan den bezala (4.3 testu-kutxa) NAF fitxategia sortzeko, analizatuko den deskribapena `.txt` batean egon beharko da. Beraz, arazorik ez sortzeko, deskribapenak kargatzen diren bakoitzean aldi baterako `.txt` fitxategi batean idatziko dira (bi erabiltzaile aldi berean erabiltzen badute sisteman arazorik ez sortzeko) eta hemendik `eustagger-lite`-ri deituko zaio NAF fitxategia sortzeko (ikusi 6.1 irudia).

`Wrapper` hau sortu da, `Med_en` sisteman terminoen informazio lexikoa JSON elementuetan gordetzen direlako. Beraz, JSON elementu hauek erabiliko dira `Med_eu` sisteman aldaketa gutxi egiteko.



*Irudia 6. 1: Wrapper-aren itxura.*

Beraz, aldi baterako fitxategia sortzeko (Hellmann, 2016) ikusi F eranskinetako F.1 testu-kutxako `terminoTemp()` funtzioa.

Aldi baterako fitxategia sortua izan dela, `import xml.etree.ElementTree as ET` erabiliko dugu NAF-etik (ikusi 4.3.2.1 azpiatala) lema, etab. lortzeko eta JSON itxura emateko deskribapenari (F.2 testu kutxako funtzioko 9. lerroa).

Hau da, `naf2json` funtzioaren bitartez (F.3 testu-kutxa) XML batetik (NAF) JSON bat sortuko dugu. Gainera, funtzio errekursibo bat sortua izan da case `array`-tik laburtzapean guztiak lortzeko modu eraginkor batean (laburtzapeanak funtzioa, F.4 testu-kutxan aurkitzen dena).

`tokenEtiketak()` deitzen den funtzioan (6.26 testu-kutxa) etiketa bakoitzari dagokion multzoa ezarriko zaio (kodearen zati bat erakutsiko da funtzioa oso luzea baita):

```
def tokenEtiketak(self, laburdura):
    ...
    etiketa = []
    #Puntuazio markak
    if 'PUNT' in laburdura:
        etiketa.insert(0, 'PUNT')
        etiketa.insert(1, laburdura)
    #Ortografia
    elif laburdura == 'HAS_MAI' or laburdura == 'DEN_MAI':
        etiketa.insert(0, 'ORT')
        etiketa.insert(1, laburdura)
    ...
    return etiketa
```

*Testu-kutxa 6. 26: etiketa bakoitzari bere multzoa esleitu.*

Beraz, hau guztia egin ondoren `Eustagger_lite` bueltatzen duen informazio guztia JSON moduan lortuko genuke. *Wrapper*-aren emaitza, 6.27 testu-kutxan ikusten duguna izango litzateke.

```
{'sentences': [[['brakizefalia', {'CharacterOffsetEnd':  
'11', 'MUG': 'MUGM', 'NamedEntityTag': 'O', 'NUM': 'NUMS',  
'POS': 'IZE', 'CharacterOffsetBegin': '0',  
'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'],  
'Lemma': 'brakizefalia', 'CASE': 'ABS', 'SubPOS': 'ARR'}]]]}
```

*Testu-kutxa 6. 27: : jsonFinala aldagaia bueltatuko duena (ikusi Feranskina)*

Behin hau eginda `TermMed_eu` zerbitzaria eta `AnaMed_eu` bezeroaren sorketarekin ekingo diogu *wrapper*-rak sortzen dituen JSON terminoetatik informazio gehigarria gehitzeko (hierarkia, SNOMED identifikatzailea, ...).

## 6.5 Med\_eu-en sorkuntza

Aurretik esan dugun bezala, `Med_eu` sistema sortzeko `Med_en`-en sistema hartu eta hainbat aldaketak egin behar izan dira. Bi fitxategietan aldaketak egin behar izan dira: `TermMed_eu` zerbitzarian (`ts_eu.py`) eta `AnaMed_eu` bezeroan (`analizatzailea_eu.py`).

### 6.5.1 TermMed\_eu:

Hemen soilik, deskribapenak, erlazioak eta kontzeptuak euskaraz izan behar direnez beste fitxategien izenak atzitu behar izango ditugu. Portuaren zenbakia aldatu behar izango dugu (kasu honetan, portua 9608 izango da) arazorik ez egoteko (6.28 testu-kutxatik beltzez agertzen den kode zatia izango ziren fitxategi honetako moldaketak).

```

def __init__(self, inp):
    m = re.match('.*\/SnomedCT_RF([12])Release_([^\_]*)([^\/]*)\/', inp)
    if m:
        fileType = "sct"+m.group(1) #
        namespace = m.group(2) #eu izango da
        version = m.group(3) #20150731 izango da
    else:
        print('fitxategiaren izena ez da ezagutu')
        exit()
    fitxC = inp+fileType+' _Concepts_Core_' +namespace+'_'+version+'.txt'
    fitxD = inp+fileType+' _Descriptions_' +namespace+'_'+version+'.txt'
    fitxR = inp+fileType+' _Relationships_Core_' +namespace+'_'+version+'.txt'
    ...

def main():
    """
    The code below starts an JSONRPC server
    """
    parser = optparse.OptionParser(usage="%prog [OPTIONS] ")
    parser.add_option('-p', '--port', default='9608',
                    help='Port to serve on (default 9608)')
    ...
    inp =
'/sc01a4/users/aelorz003/KBP/baliabideak/SnomedCT_RF1Release_eus_20150731/'
    des = HierarkiakKargatu(inp)
    ...

```

Testu-kutxa 6. 28: ts\_eu.pyfitxategia.

Behin hau eginda, AnaMed\_eu erremintan moldaketekin hasiko gara, eponimoen era egokian tratatzen.

### 6.5.2 AnaMed\_eu:

Fitxategi honen aldaketa nagusia eponimoak modu egoki batean identifikatzea izan da (Downen sindromea terminoan ez zuen Down eponimotzat identifikatzen, adibidez). Horretarako, eponimoakIdentifikatu() funtzioan eponimoak.txt fitxategitik eponimo guztiak lortuko dira. Erabiltzaileak idatzi duen hitza fitxategian aurkitzen bada, lema moldatuko dugu eponimoaren izenera. Beti eponimo luzeena lortuko dugu (ikusi 6.29 testu-kutxa):

```

Eponimoak.txt -> Down, Barré, Park, Pariknson, ...
Parkinsonen gaixotasuna -> Park      Parkinsonen gaixotasuna -> Parkinson
parkinson bueltatuko du lematzat bien artean luzeena delako

```

Testu-kutxa 6. 29: eponimo luzeenaren lorketa.

Eponimo luzeenaren kontrola egiteko izateko 6.30 testu-kutxa ikusi:

```
#izen bereziak identifikatzeko. EPONIMOAK = Down, Hopkins,...
def eponimoakIdentifikatu(tagged):
    ...
    lagLemmaEponimo = "
    for i in range(len(eponimo)):
        #park eta parkinson bezalako bereizketak egiteko
        #(formarekin konparatu). len() erabiliko da eponimo
        #luzeena .txt fitxategitik lortzeko.
        if len(lagLemmaEponimo) < len(eponimo[i].lower())
and eponimo[i].lower() in forma.lower() and forma.lower()[0:
len(eponimo[i])] == eponimo[i].lower():
    info["Lemma"] = eponimo[i].lower()
    info["NamedEntityTag"]='EPONYM'
    lagLemmaEponimo = info["Lemma"]
    aurkituta = True
    ...
```

*Testu-kutxa 6. 30: euskaraz eponimoak egoki tratatzeko kodea.*

Euskarazko bertsioan, zerbitzari bakarra erabiltzen dugunez, wrapper\_eustagger.py-ko Wrapper paketea inportatu beharko dugu. Horretarako, hurrengo hau idatzi beharko dugu soilik terminoen analisiak egiteko 6. 31 testu-kutxako analizatu() funtzioan:

```
from wrapper.wrapper_eustagger import Wrapper
...
def
analizatu(term, luzeenaBool=False, multzokatu=False, abstrakzi
oak=False):
    ...
    wr = Wrapper()
    result = wr.parse(term)
    ...
```

*Testu-kutxa 6. 31: sortutako Wrappera pakete bezala inportatu.*

result aldagaiak 6. 26 testu-kutxako informazioa bueltatuko digu. result aldagaiko sentences barruan dagoen informazio guztia nahi dugunez 6.23 testu-kutxako kodea idatziko dugu:

```
eponimoekin = eponimoakIdentifikatu(result["sentences"])
```

Hau

*Testu-kutxa 6. 32: json informazioa lortu.*

eginez, analizatzaileak 7.1 azpiataleko dagoen informazioa bueltatuko digu.



Behin hau guztia eginda, OsatEUS-en interfazea moldatzen hasiko gara erabiltzaileak sartzen duen testutik termino guztiak detektatzeko.

## 6.6 Django interfazea

Hona hemen osasunTxostenak *framework*<sup>6</sup>-ak bukaeran hartuko duen itxura (azpimarratuak izan diren fitxategiak ikasleak sortu ditu eta beltzez daudenak moldatuak izan dira):

```
osasunTxostenak
  db.sqlite3
  manage.py
  templates
    admin
      base_site.html
  txantiloiak
    init.py
    admin.py
    forma.py
    models.py
    urls.py
    views.py
    analizatzailea eu.py
    wrapper eustagger.py
    sct4text.py
    migrations
      init.py
      0001_initial.py
      ...
    static
      templatetags
        __init__.py
        txantiloiak_extras.py
      txantiloiak
        styles.css
    templates
      registration
        login.html
        login_zah.html
      txantiloiak
        espezialitatea.html
        index.html
        laburpena.html
        proba.html
      templatetags
        __init__.py
        txantiloiak_extras.py
  txostenak
    __init__.py
    settings.py
    urls.py
    wsgi.py
```

Testu-kutxa 6. 33: Django framework-a bukaeran izango duen itxura.

---

<sup>6</sup> egitura bat da, non softwareko beste proiektu bat garatu daitekeen.

6.2 irudian ikusten dena hasiera batean genuen interfazea da.

## Digestio-aparatua

<input checked="" type="checkbox"/> <b>OSPITALERATZEKO ARRAZOIA</b> <input type="text" value="as"/>
<input checked="" type="checkbox"/> <b>AURREKARIAK</b> <input type="text" value="Ez du alergiairik. Ez du ohiko tratamendurik."/>
<input checked="" type="checkbox"/> <b>EGUNGO HISTORIA</b> <input type="text"/>
<input checked="" type="checkbox"/> <b>AZTERKETA FISIKOA</b> <input type="text"/> <input type="checkbox"/> <b>Burua eta lepoa.</b> <input type="checkbox"/> <b>Biriken auskultazioa.</b> <input type="checkbox"/> <b>Sabela.</b> <input type="checkbox"/> <b>Gorputz-adarrak.</b> <input type="checkbox"/> <b>Ondeste-ukipena.</b>
<input checked="" type="checkbox"/> <b>PROBA OSAGARRIEN LABURPENA</b> <input type="text"/>

*Irudia 6. 2: Digestio-apatuaren interfazea hasieran*

Eta bukaeran lortuko dena (6.3 irudia).

The screenshot shows a web interface for 'Barne medikuntza' (Internal Medicine). It features a text area with the title 'OSPITALERATZEKO ARRAZOIA' containing a paragraph of text. Several words in the text are highlighted in yellow, and a legend on the right explains these highlights: 'Lotuta dator, asaldatau' (highlighted in green), 'biguria-egokiarekin' (highlighted in purple), 'diskurso eragabea' (highlighted in blue), 'ozenki hitz egiten du eta aldamenekoak infatzen ditu gauero hitz egin eta ez dirizelako lorik egiten ezan' (highlighted in yellow), 'esaten du kamerak jarri eziziotela' (highlighted in purple), 'osatu zer egiten duen jakiteko' (highlighted in yellow), 'inbidia omen diote' (highlighted in purple), 'bizia luzatzeko substantzia bat asmatu duelako eta lapurtu nahi diotela' (highlighted in yellow), 'Hizkera ez da batere antolatua' (highlighted in purple), 'Mesjefati dago eta ez du batere laguntzen' (highlighted in yellow), and 'Oso argal dago' (highlighted in purple). Below the text area is a red bar with the title 'AURREKARIAK' and a sub-header 'asaldatau'. A list of SNOMED terms follows, with the last one selected:

- 286582007 - Ecstatic (finding)
- 3646006 - Lamellipodium (cell structure)
- 39869006 - Alarm, device (physical object)
- 9251003 - Jumping (observable entity)
- 258141001 - Lifting (observable entity)
- 260379002 - Impaired (qualifier value)
- 18307000 - Altered (qualifier value)

Irudia 6. 3: Barne-medikuntzaren interfazea bukaeran.

OsatEUS aplikazioan egin diren aldaketa nagusiak hurrengo hauek dira: osasun-txostenak sortzeko formulario batetik, erabiltzaileak formulario horretan sartzen dituen testuaren terminoen detektatzea eta hauen informazioaren (SNOMED identifikatzailea eta dagokion FSN-ak) erakusketara pasa da. Hau guztia lortzeko, 4.5.2 azpiataleko (ikusitako azpiatal honen taulakhobeto ulertzeko) teknologia erabili da. Hainbat aldaketak egin ditugu proiektuan helburu hau lortzeko.

### Formularioa moldatzen:

Formularioaren posizioa moldatzea izango da egingo den lehenengo urratsa, hemen idazten den testua ondoren beste toki batean ikusteko (hau da, termino detektatzean lehenengo urratsa interfaze honen elementuen posizioak moldatzea izango zen).

Horretarako, `espezialitate.html` fitxategiko HTML kode hau gehitua izan da (6.34 testu kutxako beltzez agertzen den kodea). Testu-kutxetan (6.34 testu-kutxako `<textarea>` etiketak) idazten den testua beste toki batean agertuko da kurtsorea testu-kutxatik kentzerakoan (`onblur()` gertaera erabiliz hau lortuko dugu). Testu berria, sortu berri diren kutxatan agertuko da (6.38 testu-kutxako `<div>` etiketa). Hau lortzeko, id atributuak

beharko ditugu (4.5.2 azpiatala) adierazteko zein kutxetan nahi dugun testua agertzea.

```
{% for atal in espezialitatea.atala_set.all %}
  <div id=emaitzaDiv{{atal.id}} class="emaitza">
    <span id="emaitza{{ atal.id }}"></span>
  </div>
  <div class="atal">
    ...
    <textarea form="txostena" onblur =
"terminoakLortu('{{ atal.id }}')" id="testu{{ atal.id }}"
name="testu{{ atal.id }}"> {{atal.testua}}</textarea>
    ...
  </div><!--atal-->
{% endfor %}
```

*Testu-kutxa 6. 34: espezialitatea.html moldatu/gehitutako kodea (beltzez).*

Sortu berri izan diren kutxak (<div> etiketa dutenak) interfazearen eskuinaldean ezarri dira CSS erabilita. Horretarako, bi klase berri sortuak izan dira (class atributua) emaitza deiturikoa. Gainera, .atal klasetik atributu batzuk kendu ditugu jadanik beharrezkoak ez zirelako (tamaina eta posizioa aldatu denez, atributu batzuk (length bezalakoak) ez ziren beharrezkoak). Kutxan moldaketekin amaitzeko, tamaina jakin bat eman zaizkie sortu berri diren kutxari.

```

.emaitza{
    position: absolute;
    left: 53%;
    width: 45%;
    size:80%;
}

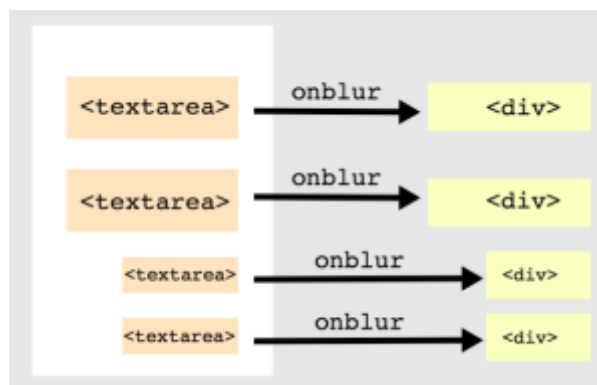
.emaitzaAzpi{
    position: absolute;
    left: 136%;
    width: 65%;
    size:80%;
}

.atal{
    position: relative;
    left: 2%;
    color:black;
    border-style: solid;
    border-width: 1px;
    size:80%;
    width: 48%;
    background-color:white;
}

```

Testu-kutxa 6. 35: espezialitatea.html-en egindako moldaketak CSS arloan.

Behin hau eginda, *JavaScript* kodea gehitu zaio `onblur()` atributuari testu-kutxetan idatzita dena kutxen barruan (`<div>` etiketa) agertzeko kurtsorea kentzerakoan. Kutxa hauek soilik ikusgarri egongo dira testua agertu behar denean (ikusi 6.4 Irudia).



Irudia 6. 4: erabiltzaileak idatzitako testua lortzeko irudi grafikoak.

Beraz, behin-behineko `jsonLortu()` funtzioa sortu dugu (6.37 testu-kutxa) testu-kutxetan dagoen testua lortzeko eta sortu berri diren kutxetan agertzeko.

```

//textarean idatzitako testua lortu kurtsorea kentzerakoan
function jsonLortu(ida) {
    var testuID = "testu".concat(ida);
    var emaitzaID = "emaitza".concat(ida);
    //textareako testua lortu
    var x = document.getElementById(testuID).value;
    var arraya = x.split(" ");
    //analizatzailea_eu.py-ri textareako testua bidali
    analizatzeko
        hitzakDetektatu(x, emaitzaID);
}

```

Testu-kutxa 6. 37: jsonLortu behin-behineko funtzioa

```

//idatzitako testua ondoko div-ean agertuko da.
function hitzakDetektatu(x, emaitzaID){
    document.getElementById(emaitzaID).innerHTML = x;
}

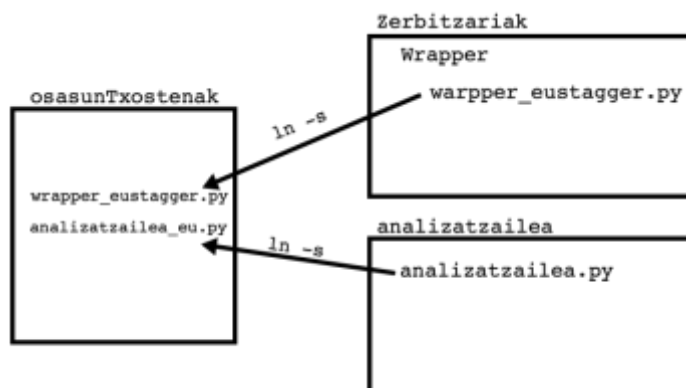
```

Testu-kutxa 6. 36: hitzakDetektatu behin-behineko funtzioa.

Hau egin ondoren erabiltzaileak idazten duena, interfazearen eskuinaldean agertuko zaio.

### AnaMed\_eu-rekin konexioa:

Testu-kutxetatik (Textarea etiketa) sartzen den testua analizatzeko AnaMed\_eu-ren beharra izango dugu. Hau egin behar denez kurtsorea testu-kutxetik kentzen den bakoitzean, jsonLortu() deituriko *JavaScript* funtzioaren barnean AnaMed\_eu bezeroari deitu behar izango diogu. Hau egin ahal izateko link birtual bat sortu beharko dugu (6.5 irudia)



Irudia 6. 5: analizatzailea\_eu eta wrapper\_eustagger.py-ren link birtualaren irudikapena.

Behin hau eginda, `jsonLortu` funtzioa honela moldatu dugu (beltzez agertzen dena):

```
//textarean idatzitako testua lortu kurtsorea
kentzerakoan
function jsonLortu(ida) {
    var testuID = "testu".concat(ida);
    var emaitzaID = "emaitza".concat(ida);
    //textareako testua lortu
    var x = document.getElementById(testuID).value;
    var arraya = x.split(" ");
    //analizatzailea_eu.py-ri textareako testua bidali
    analizatzeko
    $.ajax({
        type: "GET",
        url: "berria/analiza/",
        data: { param: x},
        dataType: "text"
    }).done(function( jsonLortu ) {
        //jsonLortu string moduan bueltatu
        hitzakDetektatu(jsonLortu, x, emaitzaID);
    });
}
```

*Testu-kutxa 6. 38: AJAX-en erabilera terminoak detektatzeko.*

Hau da, AJAX teknologia erabili dugu (Docs.python-requests.org., 2016). AJAX-ek hurrengo `url`-a bilatuko du `x` parametroa esleituz. GET motakoa esan diogu izateko, deiak ondo jasotzen dituela ikusteko Django zerbitzarian. Deia modu arrakastatsuan bete bada, `hitzaDetektatu` funtziora joko du. `jsonLortu` aldagaia interfazeko testu-kutxaren testuko hitz guztien informazio lexikoa bueltatuko du (6.38 testu-kutxako 14. lerroa). Hau da, terminoak diren hitzak SNOMED CT-ko informazioa azalduko (aurretik aipatuak izan diren, hierarkia, `sctId`, ...) da aldagai honetan. `url`-a `urls.py` deituriko fitxategian bilatuko du.

`urls.py` fitxategian hurrengo bi patroiz gehitu dira (6.39 testu-kutxa).

```
from . import views

urlpatterns = [
    ...
    url(r'^([0-9]+)/aldata/berria/analiza/$',
        views.analizatzailea, name='analiza'),
    url(r'^([0-9]+)/berria/analiza/$',
        views.analizatzailea, name='analiza'),
    ...
]
```

*Testu-kutxa 6. 39: `urls.py` fitxategia gehitutako lerroa.*



Idatzi berri izan den bi lerro hauekin (6.39 testu-kutxa) hurrengo hau esanahi da: `views.py` (6.39 testu kutxa) fitxategian, `analizatzailea` den funtzioa deitzea. Hau da, `urls.py` fitxategia `.html`-ak diren eta `views.py` arteko bitartekari soil bat izango litzateke (ikusi 6. 40 testu-kutxa eta 6.6 irudia prozesua hobeto ulertzeko).

```
1.- espezializaioa.html /berria/analiza patroia deitu.
2.- urls.py /berria/analiza patroia bilatu.
3.- urls.py views.analitzzailea funtzioa deitu.
4,. Views.py analizatzailea funtzioa exekutatu.
```

*Testu-kutxa 6. 40: AJAXetik AnaMed\_eu deitzeko prozesua.*

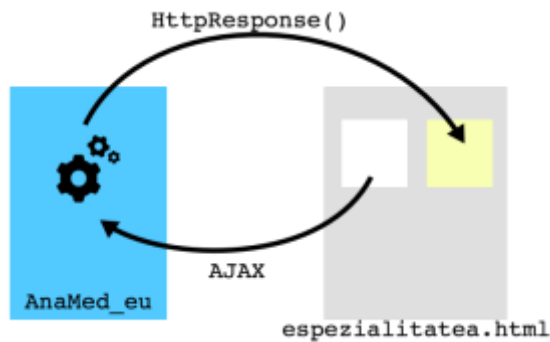
Beraz, `anaMed_eu` fitxategia atzitu ahal izateko hurrengo hau idatzi dugu `views.py`-n (6.41 testu kutxa)

```
from subprocess import Popen, PIPE
import subprocess

def analizatzailea(request, param):
    //python3 analizatzailea_eu.py "dagokion testua"
    deitu.
    subprocesLortu = subprocess.Popen(['python3',
    '/sc01a4/users/aelorz003/KBP/osasunTxostenak/txantiloak/an
    alizatzailea_eu.py', request.GET['param']], stdout=PIPE,
    close_fds=True)
    analizatu = subprocesLortu.communicate()
    analizatu = analizatu[0].decode('utf-8')
    print (analizatu)
    return HttpResponse (analizatu)
```

*Testu-kutxa 6. 41: views.py fitxategia.*

Bukatzeko, esan beharra dago `request` aldagaia aldagai finko bat dela (Django erabiltzen diren funtzio guztietan ezarri behar den parametro bat da). Behin `AnaMed_eu` sistemak interfazeko testu-kutxaren analisia egin ondoren, lortutako emaitza `espezializazioa.html` fitxategira bidaliko dugu `HttpResponse ()` funtzioaren bitartez (6.46 testu-kutxa).



Irudia 6. 6: anaMed\_eu sistemaren erantzuna Django-n lortu

Hau guztia eginda, Anaterm\_eu erramintaren bitartez interfazeaz sartua izan den testu-kutxako testu guztia analizatua izango da terminoak zein diren identifikatuz. Beraz, orain detektatuak izan diren terminoen kolore etiketatzea egingo dugu.

### Terminoak koloreztatzen:

Terminoak hierarkien arabera sailkatzen dira (DISORDER, FINDING,...). Hemeretzi hierarkia mota desberdin ditugunez, sei hierarkia mota nagusiak aukeratuko ditugu honela eta bakoitzari kolore bat esleitu zaio:

Garrantzia	Hierarkia	Etiketa kolorea	HTMLetiketa
1	DISORDER	Gorrixka	#ff8080
2	FINDING	Berdea	#99ff99
3	PROCEDURE	Urdin argia	#00ffff
4	BODYSTRUCTURE	Arroxa	#e600ac
5	PHARMPRODUCT or SUBSTANCE	Laranja	#ffaa00
6	QUALIFIER	Horia	#ffff00
7	BESTEAK	Morea / urdin iluna	#b3b3e6

Taula 6. 7: hierarkiak eta hauen lehentasunak

Termino bat anbigua izan daiteke, hau da hierarkia bateko edo gehiagoko etiketa izan dezake (gure produktuak lortzen dituen termino anbiguoak azpimarratuak egongo dira). Termino hauen etiketa kolorea 6.7 taularen arabera esleituko dira (ikusi 6.42 testu-kutxa)

Adenokartzinoma-ren hierarkia: DISORDER#BODYSTRUCTURE

Testu-kutxa 6. 42: Adenokartzinoma-ren bi hierarkiak.

Beraz, adenokartzinoma-ren etiketa kolorea gorrixka izango da, DISORDER etiketak aurreko taulan BODYSTRUCTURE etiketak baino lehentasun handiago duelako. 6.43 testu-kutxako kodea gehitu diogu hitzakDetektatu() javascript funtzioari kolorezko etiketak esleitzeko terminoari. Funtzio hau gauza asko egiten dituen eta oso luzea denez, hainbat zatitan banatu dugu hobeto azaldu ahal izateko (ikusi 6.43 testu-kutxatik 6.49 testu-kutxara).

```
// '-k ordez " idatzi
var bitartekaria = testua.replace(/'/g, '"');
// String-a JSON bihurtu
var jsonTestua = JSON.parse(bitartekaria);
// termino guztiak dagokion kolorearekin azaldu
var testuaFinal = "";
var dauka = false;
```

Testu-kutxa 6. 43: String - JSON konbertsioa.

6. 43 testu-kutxan aurretik sortu den aldagaia (jsonLortu, alegia) string formatutik JSON formatura pasako dugu. Horretarako, string-ean aurkitzen ditugun kakotx bakun guztiak (‘’) kakotx bikoitzera bihurtuko ditugu ( ‘ ” ’). Javascript-en aldagai batek JSON motakoa izateko bakarrik kakotx bikoitzak onartzen ditu balioetan (beste lengoaietan ez bezala). Hona hemen eredu bat:

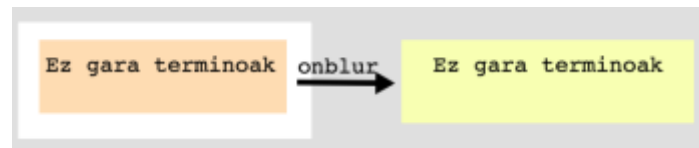
<pre>["Ospitaleratzean", <b>'PROCEDURE-0'</b>, 'ospitaleratze', <b>'sctId': '394656005'</b>, 'CASE': 'INE', 'ORT': 'HAS_MAI', 'MUG': 'MUGM', 'CharacterOffsetBegin': '0' 'NamedEntityTag': 'O', 'SubPOS': 'ARR', 'NUM': 'NUMS' 'CharacterOffsetEnd': '14', 'POS': 'IZE' 'funtzioSintaktikoak': ['@ADLG']]</pre>	<pre>{'Hierarchy': <b>"PROCEDURE-0"</b>, "Lemma": "ospitaleratze", <b>"sctId": "394656005"</b>, "CASE": "INE", "ORT": "HAS_MAI", "MUG": "MUGM", "CharacterOffsetBegin": "0", "NamedEntityTag": "O", "SubPOS": "ARR", "NUM": "NUMS", "CharacterOffsetEnd": "14", "POS": "IZE", "funzioSintaktikoak": ["@ADLG"]}</pre>
---	--

Hau egin ondoren, testuaFinal (kolorez dauden termino etiketatuek eta ez daudenak kateatzeko) eta dauka aldagai boolearra

Testu-kutxa 6. 44: 6.43 testu-kutxako jsonLortu (string) eta jsonTestua (JSON) aldagaia.

(ikusteko ea terminoa ea etiketatuta izan den) sortu ditugu aurrerago ikusiko dugun moduan. Behin hau eginda, jsonTestua aldagaian dauden terminoak identifikatzea izango da. Beraz, terminoa izateko sctId edo Hierarchy aldagaiak egon beharko dute (6. 43 testu-kutxan urdinez dauden etiketak). Kasu honetan, jsonTestua aldagaian sctId aldagaia (6. 44 testu-kutxaren

sctId-a) saiatuko gara lortzen. Hau da, lehenengo if horren barruan sartzen diren hitzak, ez dira terminoak izango (ez zaie kolorezko etiketarik ezarriko) sctId aldagaia ez dutelako (ikusi 6.45 testu-kutxa). testuaFinal aldagaian kateatuko dira, baino etiketa kolore bat ezarri gabe (6.7 irudia).

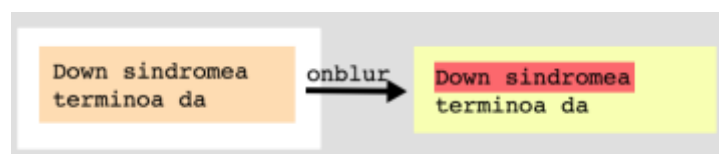


Irudia 6. 7: terminoak ez diren hitzen erakusketa interfazean.

```
for(i = 0; i < jsonTestua.length; i++){
    // hitzak ez badauka snomedId-rik ez badauka -> ez
    zaio kolorerik esleituko
    if (typeof jsonTestua[i][1].sctId == "undefined"){
        testuaFinal += jsonTestua[i][0]+ " ";
    }
    else{
        ...
    }
}
```

Testu-kutxa 6. 45: identifikatu zeintzuk ez diren terminoak.

else barruan hau egingo dugu: termino anitzak eta anbiguoak diren terminoak detektatu eta etiketa koloreak esleitu dizkiegu (6. 8 irudia).



Irudia 6. 8: termino anitzen detektatzea eta etiketa eisleipena

Termino anitzak detektatzeko termino bakoitzaren sctId-an oinarritu gara. Hau da, sctId-an “\_HAS”, “\_ERD” edo “\_BUK” bezalako atzizkiak aurkitzen baditugu, termino anitz baten parte dela esanahi du. Beraz, 6.47 testu-kutxako kodea idatzi dugu else-aren barruan.

```
Down-en: 41040004_HAS
Sindromea: 41040004_BUK
```

Testu-kutxa 6. 46: "Down-en sindromea" termino anbiguoaren identifikatzaileak.

Hau da, \_HAS atzizkia aurkitzen duenean indexOf() funtzioaren bitartez, termino anitz baten hasiera adieraziko du. Hau hasieratzerakoan j

deritzon kontagailua aktibatuko da eta hitzak kateatzen joango da termino anitza sortuz. Behin `_BUK` atzizkia aurkitutakoan termino anitzaren bukaera dela adieraziko du eta termino honen etiketatzeari jarraituko dio (kolore bat esleituz). Ikusi 6.50 testu-kutxa.

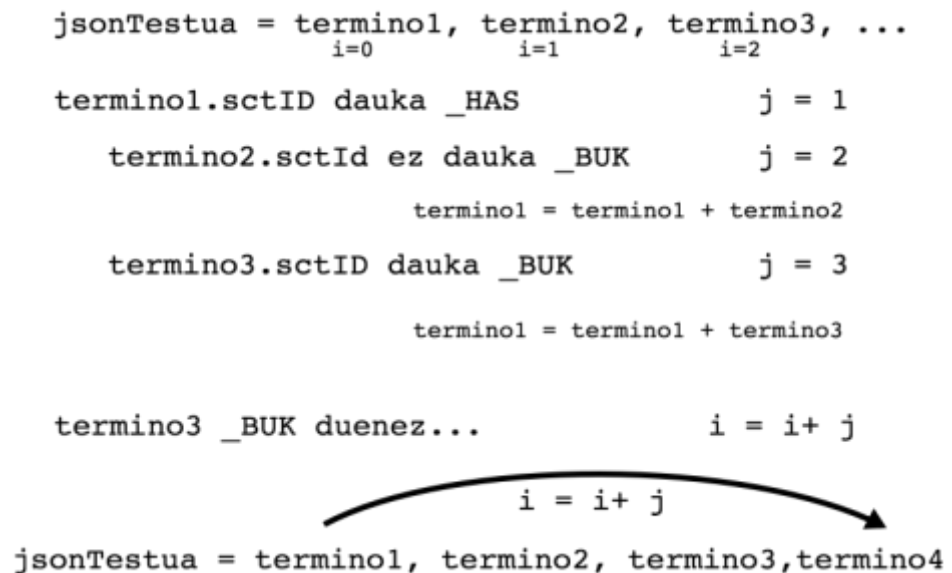
```

var j = 0;
//snomedId "HAS" badauka, terminoak konkatenatuko ditu (hitz
anitzeko terminoentzako)
if (jsonTestua[i][1].sctId.indexOf("HAS") > -1) {
    while((jsonTestua[i][1].sctId.indexOf("BUK")) <= -1) {
        if (jsonTestua[i+j][1].sctId.indexOf("BUK") > -1){
            break;
        }else{
            j++;
            jsonTestua[i][0] = jsonTestua[i][0].concat("
".concat(jsonTestua[i+j][0]));
        }
    }
}
...
if (jsonTestua[i][1].sctId.indexOf("HAS") > -1) {
    i = i + j;
}

```

Testu-kutxa 6. 47: termino anitzen eta bakunen kontrolaketa.

Eskema hobeto ulertzeko ikusi 6.9 irudia.



Irudia 6. 9:termino anitzak era egokian lortzeko funtzioaren sasikodea.

Ondoren terminoaren anbiguotasuna aztertuko dugu. Horretarako, berriro `sctId` atributuan oinarrituko gara. Hau da, aldagai honek “#”

karakterea badu hainbat SNOMED-ID kateatzen ari dela esan nahi du edo beste hitzetan, termino hori esanahi bat baino gehiago duela (ikusi 6.48 testu-kutxa)

```
function terminoAnbiguoak(terminoa, sctId){
  if(sctId.indexOf("#") > -1){
    terminoa = "<u>" + terminoa + "</u>";
  }
  return terminoa;
}
```

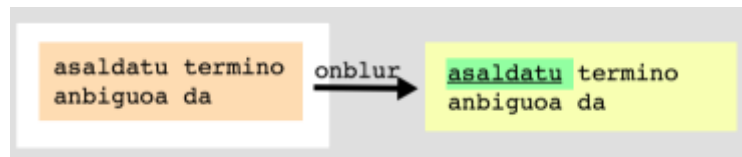
*Testu-kutxa 6. 48: terminoAnbiguoak deituriko funtzioa.*

Adibidez, Gaixotasunaren terminoa anbigua da, hiru SNOMED ID kateatzen dituelako sctId aldagaian (ikusi 6.49 testu-kutxa).

```
'sctId': '64572001#39104002#556001'
```

*Testu-kutxa 6. 49: "gaixotasunaren" SNOMED-ID-ak.*

Beraz, honela ikusiko genuke interfazean (ikusi 6.10 irudia):



*Irudia 6. 10: termino anbiguen detektatzea.*

Bukatzeko, terminoei etiketen kolorea esleitu behar zaie. Horretarako, Hierarchy atributuan begiratu beharko dugu zein hierarkia motak dauden. Hau egin ahal izateko indexOf() funtzioaren bitartez egin dugu. Funtzio honen bitartez parentesien artean idatzitakoa saiaturik da bilatzen. Hau badago, aurretik aipatu den dauka funtzio boolearra true izatera pasako da beste hierarkiak deuseztatuz (ikusi 6.50 testu-kutxa).

```

//koloreak ezarri terminoei (prioritate gehiena duen hierarkia
DISORDER >.....> gutxiena QUALIFIER
if (jsonTestua[i][1].Hierarchy.indexOf("DISORDER") > -1) {
    //kolore gorrixka = #ff8080
    testuaFinal += '<mark id =
    "'+'hitza'+'_'+emaitzaID+'_'+'i+'" style = "background-color:
    #ff8080">'+jsonTestua[i][0] + '</mark> ';
    dauka = true;
} else if ((jsonTestua[i][1].Hierarchy.indexOf("FINDING") > -1)
&& dauka == false) {
    //kolore berdexka = #99ff99
    testuaFinal += '<mark id =
    "'+'hitza'+'_'+emaitzaID+'_'+'i+'" style="background-color:
    #99ff99">'+jsonTestua[i][0] + '</mark> ';
    dauka = true;
}
//PROCEDURE, BODYSTRUCTURE, PHARMPRODUCT || SUBSTANCE eta
QUALIFIER berdina egin da
} else if (dauka == false) {
    // kolore urdin iluna-morea = #b3b3e6
    testuaFinal += '<mark onclick =
    "popUpAtera(\'hitza_' +emaitzaID+'_'+'i+\',\'+jsonTestua[i][1].
    sctId+\', \'+jsonTestua[i][1].Hierarchy+\')" id =
    "'+'hitza'+_''+emaitzaID+'_'+'i+'" style="background-color:
    #b3b3e6">'+jsonTestua[i][0] + '</mark> ';
}
dauka= false;

```

Testu-kutxa 6. 50: Terminoen etiketen kolore esleipena.

Hau guztia egin ondoren termino anitzak, bakunak eta anbiguoak direnak era egokian erakutsiko dira interfazeaz. Behin hau eginda, terminoen SNOMED Identifikatzailea eta FSN-ak (*Fully Specified Name*) *popUp* leiho batean agertuko dira termino horren gainean klik egin ondoren. Hau lortu ondoren, produkturen helburua beteko litzateke.

### PopUp leihoa agertu:

Esan den bezala, termino batean klik egin ondoren *PopUp* leiho bat agertuko da. Leiho honek hiru zatitan banatuko da, HTML, CSS eta Javascript zatietan hain zuzen ere. Leiho hau interfazearen ezker aldeko beheko aldean agertuko da beti. Leiho honek honela egongo da egituratuta (ikusi 6.12 irudia).

Hona hemen HTML kode zatia:

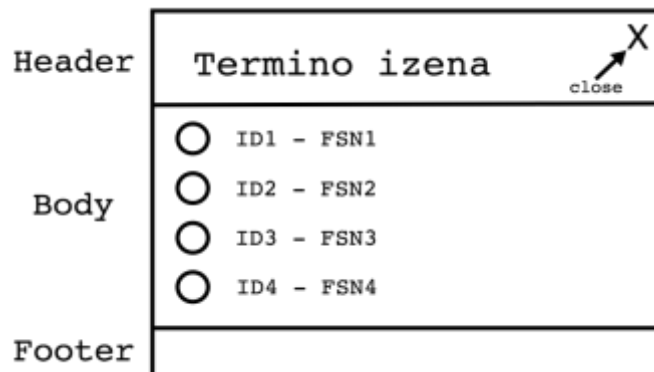
```
<div id="myModal" class="modal">

  <!-- -->
  <div class="modal-content">
    <div class="modal-header">
      <span class="close">&times</span>
      <!-- terminoaren izena hemen agertuko da -->
      <h2 id="terminoIzena"> </h2>
    </div>
    <div class="modal-body">
      <!-- radioBttnak hemen agertuko dira -->
      <blockquote id="radioButtonMultzo">
      </blockquote>
    </div>
    <!-- footerra gehitu -->
    <div class="modal-footer"> . <div>
  </div>

</div>
```

Testu-kutxa 6. 51: PopUp HTML5 kodea.

Beraz, *PopUp*-a hiru zatitan banatuko da hurrengo irudian ikus dezakegun bezala:



Irudia 6. 11: PopUp-ak hartuko duen itxura.

`<blockquote>` etiketa gehitu zaio (ikusi 6.51 testu-kutxa) *PopUp*-ari *radio button*-ak errazago ikusteko (hauen artean tarte gehiago eman, hain juntu ez ikusteko). Gainera, *footer* bat gehitu zaio, itxura erakargarriago bat emateko. Id atributuak sortuak izan dira *JavaScript* funtzioentzako eta `class` atributuak, aldiz, CSS-arentzako.



Behin popUp leihoa sortua izan dela, CSS gehitu zaio PopUp-ari animazioak gehitzeko (terminoaren gainean klik egiterakoan PopUp-a pantailaren beheko aldetik agertzeko edo X botoian klik egiterakoan hau desagertzeko, adibidez), koloreak esleitzeko eta tamaina eta posizioa finkatzeko. Hona hemen modal-ren (kutxa nagusia) propietateak. Kutxa honek, atzealdean kolorea beltza emango dio eta pantaila osoa okupatuko du. Animazioaren iraupena 0.6 segundo iraungo du (6.52 testu-kutxa).

```
.modal {
  display: none; /* defektuzki ezkutatu div-a */
  position: fixed; /* tokitik ez mugitzea (fixed) */
  z-index: 1;
  left: 0;
  top: 0;
  width: 100%; /* zabalera */
  height: 100%; /* altuera */
  overflow: auto; /* scroll baimendu behar bada */
  background-color: rgb(0,0,0); /* atzekaldeko kolorea */
  background-color: rgba(0,0,0,0.4);
  -webkit-animation-name: fadeIn; /* Fade in the background */
  -webkit-animation-duration: 0.6s; /* animazioaren iraupena */
  animation-name: fadeIn; /* animazioaren izena */
  animation-duration: 0.6s
}
```

Testu-kutxa 6. 52: modal" CSS klasea.

Honekin batera modal klaseak *webkit*<sup>7</sup> bat erabiltzen du fadeIn deiturikoa kutxaren opakutasuna %0-tik %100-ra izatera (ikusi 6.53 testu-kutxa).

```
/*animazioaren helburua: opazitatea zero-tik %100-ra igarotzea*/
@-webkit-keyframes fadeIn {
  from {opacity: 0}
  to {opacity: 1}
}

@keyframes fadeIn {
  from {opacity: 0}
  to {opacity: 1}
}
```

Testu-kutxa 6. 53: "fadeIn" keyframe-ak.

---

<sup>7</sup> Web nabigatzaile motore batekin kode askea duena.

6.54 testu-kutxan, *content*, *header*, *body* eta *footer*-rei dagozkien CSS-ak. Atzealdearen kolorea, letren kolorea etab. esleituko zaizkie:

```
/* Modal */
.modal-content {
  position: fixed;
  bottom: 0;
  background-color: white;
  width: 50%;
  -webkit-animation-name: slideIn;
  -webkit-animation-duration: 0.4s;
  animation-name: slideIn;
  animation-duration: 0.4s
}
/*modal headerra */
.modal-header {
  padding: 2px 16px;
  background-color: red;
  color: white;
}
/* modal bodya*/
.modal-body {
  padding: 1px 1px;
}
/*modal footerra */
.modal-footer{
  heigth: 3px;
  background-color: black;
}
```

Testu-kutxa 6. 54: "modal" content, header, body eta footer CSS kodeak.

Content-ari dagozkion animazioak (0.4 segundo iraunduko duena). Ikusi 6.55 testu-kutxa.

```
/* animazioa nundik nora joan den*/
@keyframes slideIn {
  from {bottom: -300px; opacity: 0}
  to {bottom: 0; opacity: 1}
}

@-webkit-keyframes slideIn {
  from {bottom: -300px; opacity: 0}
  to {bottom: 0; opacity: 1}
}
```

Testu-kutxa 6. 55: "slideIn" keyframeak.

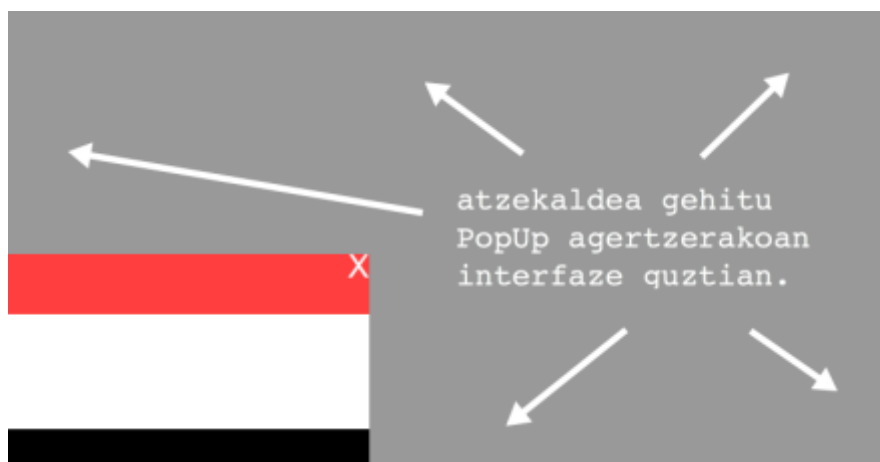
Eta X botoiak duen CSSa. Botoi hau klik egiterakoan modal kutxa desagertuko da. X botoiaren gainean xagua jartzerakoan honen kolorea txuritik beltzara pasako da (`close:focus` eta `close:hover`):

```
/* Ixteko botoia */
.close {
    color: white;
    float: right;
    font-size: 28px;
    font-weight: bold;
}

.close:hover,
.close:focus {
    color: #000;
    text-decoration: none;
    cursor: pointer;
}
```

Testu-kutxa 6. 56: X botoiaren CSS kodea.

Hau guztia idatzi ondoren klik egiterakoan termino batean PopUp hutsa lortuko genuke esleitutako kolore eta tamainarekin (6.12 irudia). Orain *javaScript* funtzioak gehituko zaizkie terminoaren izena eta dagozkien SNOMED identifikatzailea eta FSN-ak agertzeko.



Irudia 6. 12: PopUp leihoa

Egin izan den lehenengo gauza `hitzakDetektatu()` funtzioan `onclick()` atributua gehitzea izan da terminoei kolorezko etiketak esleitzerakoan (6.57 testu-kutxako beltzez agertzen kodea). *DISORDER* hierarkia soilik agertu arren, beste sei hierarkiekin egin da ere.

```

function hitzakDetektatu(testua, x, emaitzaID) {
    ...
    if (jsonTestua[i][1].Hierarchy.indexOf("DISORDER") > -1) {
        //kolore gorrixka = #ff8080
        testuaFinal += '<mark onclick =
"popUpAtera(\'hitza_'+emaitzaID+'_'+i+'\','+jsonTestua[i][1].s
ctId+'\','+jsonTestua[i][1].Hierarchy+'\')" id =
"'+hitza+'_'+emaitzaID+'_'+i+'" style="background-color:
#ff8080">'+jsonTestua[i][0] + '</mark> ';
        dauka = true;
    }
    ...
}

```

Testu-kutxa 6. 57: terminoa klik egiterakoan PopUp-a agertzeko.

Esan dugun bezala, erabiltzaileak termino batean klik egiterakoan PopUpAtera() funtzioari deituko zaio (6.58 testu-kutxa). Funtzio honek modal kutxan terminoaren hainbat informazio agertuko da (terminoaren izena header-rean, SNOMED Identifikatzailea eta dagozkien FSN-ak body-an).

```

// terminoaren gainean klik egiterakoan popUp-a azalduko da
function popUpAtera(terminoID, sctIdMultzoa,
hierarkiMultzoa) {
    var modal = document.getElementById('myModal');
    //terminoaren izena lortu
    var hitzAnbigua =
document.getElementById(terminoID).innerHTML;
    //terminoa header-ean idatzi
    document.getElementById('terminoIzena').innerHTML =
hitzAnbigua;

    //popUp-ean agertuko diren snomed eta hierarkiak
(aukeratutako terminoarena)
    var radioBttMultzoa =
lortuRadioButton(sctIdMultzoa);

document.getElementById('radioBttMultzo').innerHTML =
radioBttMultzoa;
    //modal agertu pantailan
    modal.style.display = "block";
}

```

Testu-kutxa 6. 58: PopUp-aren elementuak.

Bukatzeko, SNOMED Identifikatzaileak eta dagozkien FSN-ak *radio button*-etan agertzeko `lortuRadioButton()` funtzioa sortu egin da (ikusi 6.59 testu-kutxa).

```
function lortuRadioButton(sctMul){
    var radioMultzoak = "";
    var stcIdArray = sctMul.split("#");
    for(i = 0; i < stcIdArray.length; i++){
        // _HAS, _ERD edo _BUK ezabatu sctId-tatik
        if(stcIdArray[i].indexOf("_") > -1){
            var indexMarra =
stcIdArray[i].indexOf("_");
            stcIdArray[i] =
(stcIdArray[i]).slice(0,indexMarra);
        }
        //AJAX deitu snomed-ct idaren FSNa lortzeko.
        var fsnLortu = ajaxDeitu(stcIdArray[i]);

        //radioButtonak konkatenu. Hemen SnomedId eta fsn
agertuko dira.
        radioMultzoak += '<input type="radio"
name="anbigua" checked="checked" value="'+i+'"> '+ stcIdArray[i]+'
- '+ fsnLortu + '<br>';
    }
    return radioMultzoak;
}
```

Testu-kutxa 6. 59: *radioButtonei Id eta FSNa lortu.*

Esan dugunez bezala, SNOMED ID guztien FSN-ak lortzeko, AJAX teknologia erabili dugu. Hau egin baino lehen `_HAS`, `_BUK` edo `_ERD` bezalako atzikiak kenduko dizkiegu ID-etatik (ikusi 6.59 testu-kutxa). Behin hau eginda `ajaxDeitu()` funtzioari deituko diogu `sct4Text()` (6.65 testu-kutxa) funtziotik FSN-a lortzeko. Bi elementu hauek lortu ondoren (SNOMED ID eta dagokion FSN-a) *radio button*-ean gehitu ditugu string bezala (kateaketa erabili dugu).

### ***ID-tik FSN-a lortu:***

FSN-a lortzeko AJAX teknologia erabili da. Kasu honetan, AJAX sinkronotzat ezarri dugu (asinkronoa bere balio lehenetsia da).. Sinkronoa jarri dugu jarraitu behar duen prozesua ordenan egiteko (ikusi 6.60 testu-kutxa).

```

function ajaxDeitu(sctId){
    var idsnomed = sctId;
    var fsnLortu = $.ajax({
        type: "GET",
        url: "fsnLortu/",
        data: { param: idsnomed },
        async: false,
        dataType: "text"
    })
    return fsnLortu.responseText;
}

```

*Testu-kutxa 6. 60: AJAX teknologia.*

Beraz, aurretik egin den moduan `urls.py` eta `views.py` fitxategiak (ikusi 6.60 eta 6.61 testu-kutxak) moldatu egin ditugu `sct4text.py` fitxategiari deia egiteko ikusi (6. 40 testu-kutxaren prozedura).

`urls.py`:

```

urlpatterns = [
    url(r'^$', views.IndexView.as_view(), name='index'),
    url(r'^([0-9]+)/fsnLortu/$', views.sct4text,
        name='sct4Text'),
    ...
]

```

*Testu-kutxa 6. 61: urls.py fitxategia.*

`views.py`:

```

def sct4text(request, param):
    subprocessLortu = subprocess.Popen(['python3',
    '/sc01a4/users/aelorz003/KBP/osasunTxostenak/txantiloik/sct4text
.py', '-f', request.GET['param'], '-l', 'en'], stdout=PIPE,
close_fds=True)
    fsn = subprocessLortu.communicate()
    fsn = fsn[0].decode('utf-8')
    print(fsn)
    return HttpResponse(fsn)

```

*Testu-kutxa 6. 62: sct4text() funtzioa views.py fitxategian.*

Hau da, `sct4text` funtzioan SNOMED identifikatzaile bat idatziz (ikusi 6.61 testu-kutxa) parametrotzat dagokion FSN-a bueltatuko du (ikusi 6.62 testu-kutxa):

```
18307000 → Altered (qualifier value)
```

*Testu-kutxa 6. 63: SNOMED ID-aren FSN-a.*

`Subprocess.Popen()` funtzioaren bitartez (Rafalinux.com., 2015) `sct4text.py` fitxategiari deituko genioke (ikusi 6.61 testu-kutxa). Funtzio hau komando lerroa simulatzen duen funtzioa da. Hau da, `subprocess` funtzioaren bitartez komandoak `run` egiten ditu hainbat argumentu pasatuz(6.63 testu-kutxa).

```
python3 /sc01a4/users/aelorz003/KBP/osasunTxostenak/txantiloiak/sct4text.py -  
f "SNOEMDID" -l -en
```

*Testu-kutxa 6. 64: sct4text.py terminalean erabilita.*

Ikusi 6.65 testu-kutxa `sct4text.py` fitxategiaren kode guztia ikusi ahal izateko. Bi `termMed` zerbitzarien beharra izango dugu FSN-ak euskaraz edo ingelesez lortzeko. Momentuz FSN-ak soilik ingelesez daudenez `TermMed_en` zerbitzaria erabiliko dugu. Hauek euskaratuak daudenean, `TermMed_eu` edo `TermMed_en` zerbitzariak erabili ahal izango dira FSN-ak nahi diren hizkuntzan lortzeko. Gainera, momentuz soilik FSN-ak lortzeko funtzio hau inplementatu da baino hierarkiak edo `PreferreTerm`-ak lortzeko sortua dago ere. Hauen inplementazioa etorkizuneko lanerako uzten da (ikusi 6.65 testu-kutxa).

```

import argparse, json, jsonrpclib

class TermZerbitzaria:
    def __init__(self, port_number = 9608):
        self.server = jsonrpclib.Server("http://158.227.106.115:%d" %
port_number)
    def sct2fsn(self, sctId):
        return json.loads(self.server.sct2fsn(sctId), encoding="utf-8")

def sct2fsn(sctId, des):
    return des.sct2fsn(sctId)

if __name__ == "__main__":
    parser = argparse.ArgumentParser(description="")
    parser.add_argument('-l', "--language", help="Zein hizkuntzetan nahi
diren FSNak", required=True, type=str)
    parser.add_argument('-f', "--fsn", help="FSNa lortzeko SNOMED ID-
a", type=int)
    parser.add_argument('-p', "--prefTerm", help="PreferedTerm lortzeko
SNOMED ID idatzi", type=int)
    parser.add_argument('-i', "--hierarchy", help="hierarkia lortzeko
SNOMED ID-a", type=int)

    args = parser.parse_args()
    // idatzitako hizkien arabera portu batera edo bestera atzitu
    if args.language == "eu":
        port = 9608
    elif args.language == "en":
        port = 9604

    des = TermZerbitzaria(port)
    if args.fsn:
        print(sct2fsn(str(args.fsn), des))

    if args.prefTerm:
        print(sct2fsn(str(args.prefTerm), des))

    if args.hierarchy:
        print(sct2fsn(str(args.hierarchy), des))

```

*Testu-kutxa 6. 65: sct4text.py-ren edukia.*

Esan dugun bezala, aplikazioaren (6. 64 testu-kutxa) balio lehenetsia `-en` (ingelesez) da, FSN-ak bueltatzea ingelesera bueltatzeako, hauek momentuz ingelesez soilik bait daude. Hau da, `-eu` jarri arren FSN-ak ingelesez bueltatuko lituzke. Beraz, euskarako FSN fitxategia momentuz euskaratuta ez dagoenez, ingelesezko FSN-ak erabiliko ditugu.



Hau guztia egin ondoren, prest edukiko genuke gure produktua. Hau da, OsatEUS aplikazioa erabiltzaileak idazten duen testuaren terminoen detektatzea egingo luke (AnaMed\_eu analizatzaileari esker) eta hauen informazio baliagarria eskainiko luke. Hurrengo azpiatalean inplementazioaren zehar aurkitutako arazoei eta konponketei buruz hitz egingo da.

## 6.7 Topatutako arazoak – konponbideak:

Azpiatal honetan produktua egiteko aurkitu diren arazoak eta hauen konponbideak azalduko dira produktuaren kalitatea mantentzeko asmoarekin.

### *6.7.1 Informazio lexikoa sortzen:*

Hainbat moldaketa egin behar izan ditugu zerbitzariak jasotzen zituen terminoetan analisi lexiko egoki bat egin ahal izateko:

**TERMINO MOLDAKETAK ANALISI LEXIKOA EGIN AHAL IZATEKO EUSTAGGER\_LITE HERRAMINTAREKIN:**

sct1\_Descriptions1\_eus\_20150731.txt fitxategian zeuden hainbat termino, eustagerrek-lite-k ez zituen ondo hartzen (ezin zuen analisi lexikoa bueltatu) eta hauek moldatuak izan dira hitz hauen analisia egin ahal izateko. Moldaketa hauek egin ondoren (ikusi 6.71 testu-kutxa), analisisiak\_eu.json fitxategian gehitu ahal izan dira arazorik gabe.

Moldatuak izan diren termino gehienak hurrengo testu kutxan ikus ditzakegunak dira (esne-hortz, kobalt-57/58, 8-aminokuinolina, ...) eta terminoei aldaketa txikiak egin zaizkie (‘-‘ karakterea hutsunearengatik ordezkatu, antzeko esanahia duen karakterearengatik ordezkatu...). Ondoren, analisi egokia egitea lortu da.

```

if d.term[-1:] == "-":
    d.term = d.term.strip("-")
if d.term == "esne-hortz":
    d.term = "esne hortz"
if d.term == "udp-n-azetilglukosamina-lisosomal-entzime-n-azetilglukosaminefosfotransferase":
    d.term = d.term.replace("-", " ")
if d.term == "kobalt-57/58":
    d.term = "kobalt 57/58"
if d.term == "8-aminokuinolina":
    d.term = "8 aminokuinolina"
if d.term == "18-hidroxikortisona":
    d.term = "18 hidroxikortisona"
...
if d.term == "ø":
    d.term = "0"
if d.term == "1-xilulokinase":
    d.term = "1 xilulokinase"

```

*Testu-kutxa 6. 66: Moldatutako hitzak.*

RF1 fitxategitik terminoen moldaketak egin izan dira ere, hemendik aldaketa egiteko askoz errazagoa baitzen.

### HITZ ZEHATZEN EZABATZEA MOLDAKETAK EZIN ZIRELAKO EGIN:

Hainbat hitz ezabatuak izan dira, moldaketa ugari egin ondoren Eustagger-litek analizatzaileak akatsak ematen jarraitzen zuelako. Deuseztatuak izan diren hitz gehienak <sup>183m</sup>osmio edo <sup>198m</sup>thalio bezalako egitura zeukaten, beraz `d.term[0] == "^"` idatzi dugu <sup>^</sup> karakterez hasten zirenen terminoen analisia ez egiteko. Hona hemen baztertuak izan diren hitzak analisi hutsa bueltatuz:

```

errrenio[186re]
^ karakterearekin hasten diren terminoak
ñimiño
...

```

*Testu-kutxa 6. 67: Ezabatutako terminoak.*

### 6.7.2 Terminoak txostenetatik detektatzen:

Ikasleak C Eranskinetako lehenengo bi txostenak irakurri ondoren 111 termino detektatu ditu (51 lehenengoan eta 60 bigarreanean, kolorez azpimarratuak agertzen direnak). Terminoak hiru taldetan sailkatuak izan dira (7.1 azpiatalean agertzen den bezala). Bi osasun-txosten hauek irakurriak izan dira Med\_eu sistemaren funtzionamendua aztertzeko eta 3.4 azpiatalean finkatu den kalitate mahaira betetzeko. Egin diren proba guztiak 7.1 azpiatalean ikus ditzazkegu:

- Lehenengo taldean lehenengo saiakeran emaitza egokiak bueltatu dituzten terminoak agertzen dira. `sct1_Descriptions1_eus_20150731.txt`<sup>8</sup> fitxategiko informazioa lortu izan da lema era egoki tratatua izan delako eta fitxategian agertzen zelako. Termino hauek horiz azpimarratuak daude C Eranskinetako lehenengo bi osasun-txostenetan.
- Bigarren taldean `sct1_Descriptions1_eus_20150731.txt` fitxategian aurkitzen ez ziren lema gehitu dira, terminoaren analisia egiterakoan ez zuelako `sctId` aldagaia bueltatzen. Emaitza hauek modu egokian bueltatu ahal izateko, gaztelaniazko termino baliokidea lortu behar izan dugu SNOMED-CT *browser*-retik eta haren ID-a fitxategian gehitu, besteak beste. Berdez azpimarratua agertuko dira C Eranskinetako lehenengo bi osasun-txostenetan.
- Hirugarren taldean hasiera batean `eustagger_lite`-k gaizki bueltatu dituen lema terminoen analisia daude. Lema zuzentzeko, `eustagger_lite` moldatu behar izan da eta behin hau eginda, fitxategian aurkitzen ez bazen lema hori bigarren taldeko konponbidea jarraitu beharko dugu. C Eranskinetako lehenengo bi osasun-txostenetan urdinez azpimarratuak agertuko dira.

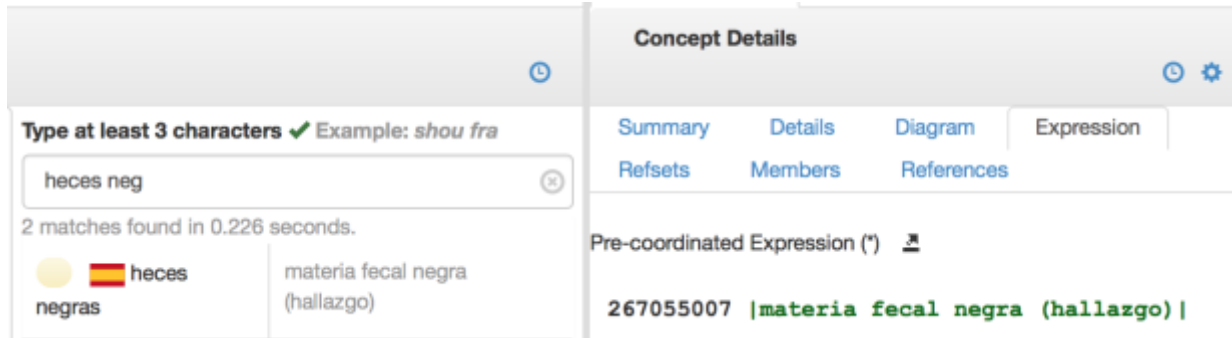
Hau da, lehenengo saiakeran ikasleak detektatu dituen termino guztietatik (111 termino) Med\_eu sistemak soilik 63 detektatu ditu. Orain bigarren eta hiru taldeko arazoei nola aurre egin azalduko ditugu.

---

<sup>8</sup> Euskarazko terminoak eta dagozkien SNOMED IDak fitxategi honetan aurkitzen dira. Med\_eu sistemak fitxategi honetatik baliatuko da jakiteko hitz bat terminoa den ala ez.

## BIGARREN TALDEKO KONPONBIDEA:

Konponketa hauek egin ahal izateko falta den termino bakoitzaren ID-ak bilatu behar izan dira. Horretarako ID-ak C eranskinetan agertzen diren gaztelaniazko txostenetan begiratu behar izango ditugu SNOMEC-CT *browser*-rean. Hau da eginkari beltzak-en gaztelaniazko terminoa materia fecal negra izango zen (ikusi 6.11 irudia).



The screenshot shows the SNOMED-CT browser interface. On the left, a search bar contains 'heces neg' and shows two matches: 'heces negras' (Spanish) and 'materia fecal negra (hallazgo)'. On the right, the 'Concept Details' panel is open for 'materia fecal negra (hallazgo)' with ID 267055007. The pre-coordinated expression is shown as '|materia fecal negra (hallazgo)|'.

Irudia 6. 13: "heces negras" terminoaren ID-aren bilaketa.

Behin ID-a lortuta, sct1\_Descriptions1\_eus\_20150731.txt RF1 fitxategian gehitu beharko genuke (6. 67 testu kutxa) eta termino horren analisia egiterakoan behar dugun informazioa lortuko genuke. Ikusten den bezala laugarren zutabean "eginkari beltzak"-en lema gehitua izan da lemen arabera bilaketak egingo direlako ("eginkari beltzak"-en lema "eginkari beltz" da):

11111167	0	267055007	eginkari beltz	0	1	eu
----------	---	-----------	----------------	---	---	----

Testu-kutxa 6. 68: RF1 fitxategian "eginkari beltz" lema gehitu.

Behin hau eginda, anaMed\_eu sistemak 37 termino berri identifikatuko lituzke C eranskinetako txostenetan.

## HIRUGARREN TALDEKO KONPONBIDEA:

Lemak gaizki identifikatzen dituzten hitzek ondo identifikatzeko eustagger\_lite erreminta moldatu behar izan da. Hau ez du ikasleak egin (Olatz Perez de Viñaspre Garraldak egin du) beraz, ezin da azaldu zein aldaketak egin diren. Behin moldaketak egin direla, ikasleak, badaezpada, bigarren ataleko prozedura errepikatu du.

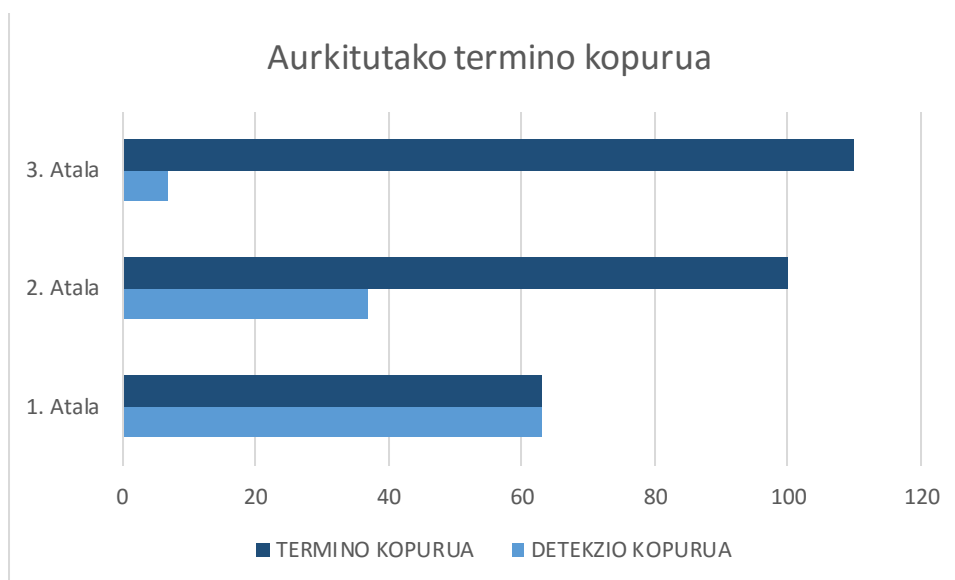
Aldaketak egin aurretik eskizofrenia paranoidearen-en lema eskizofrenia paranoidearen zen eta ondoren eskizofrenia paranoidea izatera bihurtu da.

Moldaketa hau egin ondoren 10 termino berri detektatu ditu anaMed\_eu-k.

Beraz, 111 terminotik 110 detektatzeko gai da orain. Bi txostenen %99,09-ko terminoak detektatzen ditu. Hona hemen taula eta grafikoa:

Terminoak	Detektatzea	
	Kopurua	Ehuenekoa (%)
Lehenengo atala	63	56,76
Bigarren atala	100 (63+37)	90,09
Hirugarren atala	110 (63+37+10)	99,09

Taula 6. 8: aurkitutako termino kopurua



Grafiko 6. 3: aurkitutako terminoen kopurua

Falta den terminoa (*Zyprexa Velotab*) ez da SNOMED-CT *browser*-rean aurkitu sendagai europarra delako. *Browser* honetan soilik sendagai komertzial estatu batuarrek agertzen dira edo, bestela, beste herrialdeetako sendagai arrakastatsuenak.

### 6.7.3 OsatEUS-en karaktere zehatzen ez deusaztaketa

Esan den bezala, OsatEUS aplikazioa terminoak detektatu ahal izateko, hainbat erreminta erabiltzen ditu (eustagger-lite eta SNOMED-CT bezalakoak). Eustaggerrek euskarazko esaldien analisia egiten duenez, hainbat karaktere baztertutak daude ñ edo azentuak dituzten hitzek bezalakoak.

Beraz, erabiltzaileak letra hauek idazterakoan arazorik ez sortzeko hurrengo patroia sortu du soilik karaktere jakin batzuk onartzeko, zifrak, eta eustaggerrek onartzen dituen karaktere erabilgarri batzuk:

```
var regex = /^[a-zA-Z0-9.,:;+~^? /\()]*$/;
```

Testu Kutxa 5. 3: erabiltzaileak idatzi ditzazkeen karaktereak

Hurrengo taulan Eustagger-lite-k onartzen ez dituen beste karaktereak agertzen dira:

Hona hemen Eustagger-lite erremintak onartzen ez dituen karaktereak:

Karakterea	Karakterea	Karakterea	Karaterea	Karakterea
º	'	-	Ñ	]
ª	¨	`	ñ	Ç
<	>	\$	[	ç

Taula 6. 9: eustagger-litek onartzen ez dituen karaktereak

Beraz, baldintza bat ezarriz karaktereen kontrol egokia egingo da. Sartutako karakterea ez bada egokia mezu bat agertuko zaio erabiltzaileari ohartaraziz sartutako karaktere bat ez dela onargarria:

```
if (regex.test(document.getElementById(testuID).value)) {  
...  
}else{  
    alert("Erabilitako teknologiaren ondorioz, OsatEUS aplikazioa ez du ñ edo º bezalako karakterik onartzen. Mesedez, karaktere hauek ezabatu.");  
}
```

Testu-kutxa 6. 69: Karaktereeen kontrola.

Behin produktua sortuta eta arazoei aurre eginda, probak eta ebaluazioak egin ditugu. Hauek hurrengo atalean ikus daitezke.

# 7

---

## Probak eta ebaluazioa

Inplementazio egoki bat egiten ari dela konprobatzeko hainbat proba (7.1 azpiatala) mota desberdin egin izan dira produktuan. Funtzionamendu egokia lortzen ari dela ikusteko eta finkatu izan den kalitate maila baieztatzeko egin dira. Behin espero ziren emaitzak lortu direla, proiektuaren inguruan zerikusirik ez zuten bi lagunek ebaluazioa egin dute. Ebaluazio (7.2 azpiatala) honen helburua ea terminoei ezarri zazizkien etiketa koloreak egokian ziren aztertzea izan da.

### 7.1 Probak

Proba hauek hainbat zatitan bereizten dira hurrengo orrietan ikusiko den bezala:

## Ingeles terminoen informazio lexikoa lortzen duen fitxategi txiki bat sortu:

Proba hau egin da, `.json` (terminoen informazio lexikoa gordetzen duen fitxategia) fitxategiak ondo sortzen zirela ikusteko. Fitxategi hau ez da memorian agertuko (oso handia delako), baizik eta baliabideak karpetan aurkituko da `inforLexikoa_en_proba.json` izenarekin.

## Eponimoen tratamendua:

Aurretik esan den bezala euskararen morfologia bereziki tratatu behar da. Horretarako kontuan eduki behar dira bereziki atzizkiak (-ren, -en, -ekin, ...) eta nola idatzita dauden hitzak (letra larriz, gidoitxoarekin, ...). Hona hemen Down sindromea eta Parkinson gaixotasunaren analisi posible batzuk era egokian analizatuak:

### IN: Down-en sindromea

```
OUT: [['Down-en', {'MOR': 'ZERO', 'Hierarchy': 'DISORDER-0_HAS', 'Lemma': 'down', 'NUM': 'NUMP', 'CharacterOffsetBegin': '0', 'CharacterOffsetEnd': '6', 'POS': 'IZE', 'CASE': 'GEN', 'ORT': 'HAS_MAI', 'funtzioSintaktikoak': ['@<IZLG', '@IZLG>'], 'NamedEntityTag': 'EPONYM', 'MUG': 'MUGM', 'sctId': '41040004_HAS', 'SubPOS': 'ARR'}], ['sindromea', {'Hierarchy': 'DISORDER-0_BUK', 'Lemma': 'sindrome', 'NUM': 'NUMS', 'CharacterOffsetBegin': '8', 'CharacterOffsetEnd': '16', 'POS': 'IZE', 'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'], 'CASE': 'ABS', 'NamedEntityTag': 'O', 'MUG': 'MUGM', 'sctId': '41040004_BUK', 'SubPOS': 'ARR', 'BIZ': 'BIZ-' }]]
```

### IN: Downen sindromea

```
OUT: [['Downen', {'CharacterOffsetEnd': '5', 'CharacterOffsetBegin': '0', 'MOR': 'ZERO', 'ORT': 'HAS_MAI', 'PLU': 'PLU-', 'Hierarchy': 'DISORDER-0_HAS', 'NamedEntityTag': 'EPONYM', 'Lemma': 'down', 'funtzioSintaktikoak': ['@KM>'], 'POS': 'IZE', 'sctId': '41040004_HAS', 'SubPOS': 'IZB'}], ['sindromea', {'CharacterOffsetEnd': '15', 'CharacterOffsetBegin': '7', 'Hierarchy': 'DISORDER-0_BUK', 'CASE': 'ABS', 'MUG': 'MUGM', 'NamedEntityTag': 'O', 'BIZ': 'BIZ-', 'Lemma': 'sindrome', 'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'], 'NUM': 'NUMS', 'sctId': '41040004_BUK', 'POS': 'IZE', 'SubPOS': 'ARR' }]]
```



**IN: Downen sindromea**

**OUT:** [['Down', {'sctId': '41040004\_HAS', 'CharacterOffsetEnd': '3', 'ORT': 'HAS\_MAI', 'Lemma': 'down', 'CharacterOffsetBegin': '0', 'Hierarchy': 'DISORDER-0\_HAS', 'POS': 'IZE', 'MOR': 'ZERO', 'NamedEntityTag': 'EPONYM', 'SubPOS': 'IZB', 'funtzioSintaktikoak': ['@KM>'], 'PLU': 'PLU-'}], ['sindromea', {'sctId': '41040004\_BUK', 'CharacterOffsetEnd': '13', 'Lemma': 'sindrome', 'NUM': 'NUMS', 'BIZ': 'BIZ-', 'CharacterOffsetBegin': '5', 'Hierarchy': 'DISORDER-0\_BUK', 'POS': 'IZE', 'CASE': 'ABS', 'MUG': 'MUGM', 'SubPOS': 'ARR', 'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'], 'NamedEntityTag': 'O'}]]

**IN: Parkinson-en gaixotasuna**

**OUT:** [['Parkinson-en', {'CharacterOffsetBegin': '0', 'MUG': 'MUGM', 'NamedEntityTag': 'EPONYM', 'SubPOS': 'ARR', 'sctId': '49049000\_HAS', 'funtzioSintaktikoak': ['@<IZLG', '@IZLG>'], 'CASE': 'GEN', 'Hierarchy': 'DISORDER-0\_HAS', 'POS': 'IZE', 'MOR': 'ZERO', 'NUM': 'NUMP', 'Lemma': 'parkinson', 'ORT': 'HAS\_MAI', 'CharacterOffsetEnd': '11'}], ['gaixotasuna', {'CharacterOffsetBegin': '13', 'MUG': 'MUGM', 'NamedEntityTag': 'O', 'sctId': '49049000\_BUK', 'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'], 'CASE': 'ABS', 'Hierarchy': 'DISORDER-0\_BUK', 'POS': 'IZE', 'NUM': 'NUMS', 'Lemma': 'gaixotasun', 'SubPOS': 'ARR', 'CharacterOffsetEnd': '23'}]]

**IN: Parkinsonen gaixotasuna**

**OUT:** [['Parkinsonen', {'PLU': 'PLU-', 'SubPOS': 'IZB', 'sctId': '49049000\_HAS', 'Lemma': 'parkinson', 'POS': 'IZE', 'ORT': 'HAS\_MAI', 'CharacterOffsetBegin': '0', 'Hierarchy': 'DISORDER-0\_HAS', 'CharacterOffsetEnd': '10', 'MOR': 'ZERO', 'NamedEntityTag': 'EPONYM', 'funtzioSintaktikoak': ['@KM>']}], ['gaixotasuna', {'SubPOS': 'ARR', 'sctId': '49049000\_BUK', 'Lemma': 'gaixotasun', 'CASE': 'ABS', 'CharacterOffsetBegin': '12', 'Hierarchy': 'DISORDER-0\_BUK', 'CharacterOffsetEnd': '22', 'NamedEntityTag': 'O', 'NUM': 'NUMS', 'funtzioSintaktikoak': ['@OBJ', '@PRED', '@SUBJ'], 'POS': 'IZE', 'MUG': 'MUGM'}]]

**IN: Parkinsonen gaixotasun**

**OUT:** [['Parkinsonen', {'POS': 'IZE', 'Lemma': 'parkinson', 'CharacterOffsetEnd': '10', 'sctId': '49049000\_HAS', 'MOR': 'ZERO', 'CharacterOffsetBegin': '0', 'funtzioSintaktikoak': ['@KM>'], 'PLU': 'PLU-', 'Hierarchy': 'DISORDER-0\_HAS', 'NamedEntityTag': 'EPONYM', 'ORT': 'HAS\_MAI', 'SubPOS': 'IZB'}], ['gaixotasun', {'POS': 'IZE', 'funtzioSintaktikoak': ['@KM>'], 'Hierarchy': 'DISORDER-0\_BUK', 'CharacterOffsetEnd': '21', 'SubPOS': 'ARR', 'sctId': '49049000\_BUK', 'Lemma': 'gaixotasun', 'NamedEntityTag': 'O', 'CharacterOffsetBegin': '12', 'MOR': 'ZERO'}]]

Ikusi dezakegunez, lema egokia bueltatzen du eta termino desberdinak sartu arren, sctId eta hierarkia berdinak bueltatzen ditu. Hau gertatzen da, lema berdina dutelako eta terminoen bilaketa beraien lemen arabera egiten delako.

## Osasun-txostenen probak:

Ikusi C eranskineko lehenengo bi osasun-txostenak (Alta-txosten Donostia Unibertsitateko Ospitala, V. edizioa). Ikasleak bi osasun-txosten hauetatik detektatu dituen terminoak agertzen dira hainbat tauletan. **OUT** emaitzatik bakarrik hierarkia (*Hierarchy*), sctId (SNOMED CT ID) eta lema (*Lemma*) aldagaiak agertuko dira elementu esanguratsuenak direlako.

Aurretik esan den bezala, taula hauek hiru atal desberdinetaz agertzen dira (ikusi 6.7.2 azpiatala).

### LEHENENGO TALDEA:

#### 1.txostena:

Terminoak	Lema	sctId	Hierarkia
Anemia	Anemia	271737000	DISORDER-0
Alergiarik	Alergia	609328004#408439002	DISORDER#QUALIFIER-0
Ohiko	Ohiko	70232002#40451002#6 0132005#510009#2639 01007#72906007	QUALIFIER-0
Tratamendurik	Tratamendu	119270007#276239002	PROCEDURE-0
Ahultasuna	Ahultasun	13791008#248279007# 397488002	FINDING#QUALIFIER-0
Odol-analisia	Odol-analisi	301838008	SITUATION-0
TA	TA	45215009	SUBSTANCE-0
Hg	Hg	45262002	SUBSTANCE-0
Tenperatura	Tenperatura	703421000	OBSERVABLE-0
BM	BM	419192003	QUALIFIER-0
Bihotz-maiztasuna	Bihotz-maiztasun	364075005	OBSERVABLE-0
Hidratazioa	Hidratazio	405006006	OBSERVABLE-0
Larruazal	Larruazal	181469002#39937001#2 58340006	BODYSTRUCTURE-0
Edemarik	Edema	267038008#79654002#4 23666004	FINDING#BODYSTRUCTURE#OBSERVABLE-0

Hestegorriak	Hestegorri	181245004#32849002	BODYSTRUCTURE-0
Mukosa	Mukosa	2739003#414781009#361693009#278867007	BODYSTRUCTURE-0
Lesiorik	Lesio	417163006#19130008#52988006	DISORDER#BODYSTRUCTURE-0
Piloro	Piloro	362143006#280119005	BODYSTRUCTURE-0
Ondeste-ukipena	Ondeste-ukipen	410006001	PROCEDURE-0
Sigma	Sigma	362166007#60184004#421269006	BODYSTRUCTURE#QUALIFIER-0
Itsua	Itsu	181256004#32713005#80175008	BODYSTRUCTURE#QUALIFIER-0
Biopsia	Biopsi	86273004#129314006	PROCEDURE#QUALIFIER-0
Gibela	gibel	92186001#92154006#424263008#181268008#10200004	DISORDER#BODYSTRUCTURE-0
Barea	bare	78961009#181279003	BODYSTRUCTURE-0
Pankrea	Pankrea	276851006#235963004#92264007#181277001#15776009	DISORDER#BODYSTRUCTURE-0
Behazuna	Behazun	181269000#28231008#70150004	BODYSTRUCTURE#SUBSTANCE-0
Adenokartzinoma	Adenokartzinoma	443961001#35917007	DISORDER#BODYSTRUCTURE-0
Dieta	dieta	41829006#182922004	FINDING#PROCEDURE-0
Hb	Hb	38082009#38082009	SUBSTANCE-0
Burdina	burdina	30729008#3829006	PHARMPRODUCT#SUBSTANCE-0

Taula 7. 1: Lenenengo txostenean lehenengo saiakeran detektatutako terminoak.

## 2.txostena:

Terminoak	Lema	sctId	Hierarkia
Bararrik	Bakar	267076002#420295001#50607009	FINDING#QUALIFIER-0
Agitazio	agitazio	24199005	FINDING-0
Asaldura	Asaldura	64572001#47544006	DISORDER#FINDING-0
Erretzailea	Erretzaile	77176002	FINDING-0
Hipertentsiorik	Hipertentsio	38341003	FINDING-0
Ohiko	ohiko	70232002#40451002#60132005#510009#263901007#72906007	QUALIFIER-0
Tratamendua	Tratamendu	119270007#276239002	PROCEDURE-1
Azal	Azal	181469002#39937001#258340006#228222006#706521006#706517004#23888001#257768005#410679008	BODYSTRUCTURE#OBJECT#PHARMPRODUCT#QUALIFIER#SUBSTANCE-0

		#421812003#26283006#118590008 #226740002#4962008#387128007	
Mukosetako	Mukosa	2739003#414781009#361693009#278867007	BODYSTRUCTURE-0
Perkusioa	Perkusio	75180006#418270002	PROCEDURE-0
Edemarik	Edema	267038008#79654002#423666004	FINDING#BODYSTRUCTURE#OBSERVABLE-0
Oso	Oso	255511005#4757001#11163003#255619001#60132005#260358002#264883001#255594003#255503000#260232004	QUALIFIER-0
Argal	Argal	242905004#255510006#18043004#13880007	EVENT#QUALIFIER-1
Glukosa	Glukosa	111138002#67079006	PHARMPRODUCT#SUBSTANCE-0
loiak	ioi	48006008	SUBSTANCE-0
Hemoglobina	Hemoglobina	38082009	SUBSTANCE-0
Leukozitoak	Leukozito	52501007	BODYSTRUCTURE-0
Plaketak	Plaketa	16378004	BODYSTRUCTURE-0
Koagulazioan	Koagulazio	250327000#119266004	FINDING#PROCEDURE-0
Eldarnioak	eldarnio	2776000#71038003#2073000	DISORDER#FINDING-0
Haluzinazioak	haluzinazio	7011001	FINDING-0

Taula 7. 2: Bigarren txostenean lehenengo saiakeran detektatutako terminoak

## BIGARREN TALDEA:

### 1.txostena

Terminoak	Lema	sctld	Hierarkia
Kontziente	Kontziente	271591004	FINDING-0
Orientatua	Orientatu	247663003	FINDING-0
Zeinu	Zeinu	72670004	FINDING-0'
Biguna	Bigun	26334009#264938003#37378003	OBJECT#QUALIFIER-0
Hanturarik	Hanturarik	65124004#23583003#257552002	FINDING#BODYSTRUCTURE#QUALIFIER-0
Mesfidati	Mesfidati	163614007	FINDING-0
Haluzinazioak	Halunazio	7011001	FINDING-0
Diabeteseik	Diabetesa	73211009	DISORDER-0
Bakarrik bizi	Bakar bizi	105529008	FINDING-0
Ez da garbitzen	Ez izan garbitu	284789003	FINDING-0

Talde sanitarioa	Talde sanitario	303607000	OBJECT-0
Alergia ezagunik	Alergia ezagun	160244002	SITUATION-0
peritoneoko narritadurarik	Peritoneo narritadura	281639001	FINDING-0
Eskuineko giltzurrunaren	Eskuin giltzurrun	9846003#362208000	BODYSTRUCTURE
Tronbosi-zeinurik	tronbosi-zeinu	439127006	DISORDER
Eskizofrenia paranoidea	Eskizofrenia paranoidea	191554003#31658008#64905009#63181006#191555002#79866005	DISORDER-0
ingurgitazio jugularrik	Ingurgitazio jugular	271653008	FINDING-0

Taula 7. 3: Lenenengo txostenean bigarren saiakeran detektatutako terminoak

## 2.txostena

Terminoak	Lema	sctId	Hierarkia
Kontziente	Kontziente	271591004	FINDING-0
Orientatua	Orientatu	247663003	FINDING-0
Neoplasia	Neoplasia	55342001#108369006	DISORDER#BODYSTRUCTURE-0
Gastroskopia	Gastroskopia	13219008	OBJECT-0
Eginkariak beltzak	Eginkari beltz	267055007	FINDING-0
Haztapenean ez du minik	Haztapan ez ukan min	300820003	QUALIFIER-0
Heste-soinu	Heste-soinu	249587001	OBSERVABLE
Erritmo sinusal normala	Erritmo sinusal normal	64730000	FINDING-0
Anestesia topikoa	Anestesia topiko	386761002	PROCEDURE-0
Uzkiaren ikuskapena	Uzki ikuskapen	274295008	PROCEDURE-0
Ondeste normala	Ondeste normal	300313001	FINDING-0
Beheranzko kolon	Behe kolon	32622004#362165006	BODYSTRUCTURE-0
Goranzko kolon	Gora kolon	362162009#9040008	BODYSTRUCTURE-0
Itsuko neoplasia	Itsua neoplasia	363350007	DISORDER-0

Heste itsuko neoplasia	Heste itsu neoplasia	126839008	DISORDER-0
Behazunbidea	Behazunbidea	361354009	BODYSTRUCTURE

Taula 7. 4: Bigarren txostenean bigarren saiakeran detektatutako terminoak.

## HIRUGARREN TALDEA:

### 1. Txostena:

Terminoak	Lortutako lema hasieran	Lortutako lema ondoren	sctId	Hierarkia
Eskizofrenia paranoidearen	Eskizofrenia paranoidearen	Eskizofrenia paranoide	64905009	DISORDER-0
Adenopatiarik	Adenopatiarik	Adenopatia	30746006# 181756000	DISORDER# BODYSTRUCTURE
Meningeorik	Meningeorik	meningeo	70784009#12310 04	DISORDER#BODYSTRUCTURE-0

Taula 7. 5: Lenenengo txostenean hirugarren saiakeran detektatutako terminoak

### 2. Txostena:

Terminoak	Lema Hasieran	Lema Ondoren	sctId	Hierarkia
Sabeleko minik	Sabeleko min	Sabel min	21522001	FINDING-0
Eupneikoa	Eupneikoa	Eupneiko	22803001	FINDING-0
Pultsu perfiferikoa	Pultsu periferikoa	Pultsu periferiko	54718008	OBSERVABLE-0
Batez besteko bolumen korpuskularra	Batez beste bolumen korpuskularra	Batez beste bolumen korpuskular	165455001	FINDING-0
Doudeno-erraboila	Doudeno-erraboila	Doudeno-erraboil	362149005# 9999005	BODYSTRUCTURE-0
Zeharkako kolon	Zeharkako kolon	Zeharka kolon	485005# 362163004	BODYSTRUCTURE-0
Anemia ferropenikoa	Anemia ferropenikoa	Anemia ferropeniko	87522002	DISORDER-0

Taula 7. 6: Bigarren txostenean hirugarren saiakeran detektatutako terminoak

Behin proba hauek eginda, Django *framework*-ean probak egiten hasiko gara.

## Terminoak era koherentean daudela etiketatuak baieztatu:

Moldatua izan den interfazeko testu-kutxa hurrengo hau idatzi dugu:

*Ospitaleratzean medikazio antipsikotikoa jarri zaio. Solairuan bi aldiz asaldatu da. Tratamendua hasi eta hamar egunera mesfidantza desagertu zaio, baita eldarnioak eta haluzinazioak ere. Ez du gaixotasunaren kontzientziarik, eta botika hartzea gainbegiratu behar zaio, ezkututzen saiatzen baita, beraz ospitaleratzea luzatzea erabaki da egonaldi ertaineko ospitale batean. San Juan de Dios Ospitale Psikiatrikora lekualdatu da lau hilabeterako.*

Eta hauxe bueltatu digu (lehengo bi lerroetako probak egingo ditugu nahiko errepikakorra baita prozesua):

Ospitaleratzean medikazio antipsikotikoa jarri zaio . Solairuan bi aldiz asaldatu da .  
Tratamendua hasi eta hamar egunera mesfidantza desagertu zaio , baita eldarnioak eta

*Irdia 7. 1: Terminoetako etiketa moten esleipena.*

Behin hau jasota hurrengo json emaitza jaso du AnaMed\_eu analizatzaileak. Emaitzak taula batean bistaratu izango da terminoa, lema, hierarkiak (eta honen koloreak) eta sctId-ak argiago ikusteko. Lehenetsuna duen hierarkia beltzez agertuko da. Gogora dezagun marra batez azpimarratuaz dauden terminoak termino anbiguoak direla, hau da, sctId eta hierarki bat edo gehiago dituztela.

Terminoak	Lema	sctId	Hierarkia	Kolorea
Ospitaleratzean	ospitaleratze	394656005	PROCEDURE-0	URDIN ARGIA
Medikazio	medikazio	394829006	SITUATION-1	MOREA
Antipsikotikoa	antipsikotikoa	10784006	PHARMPRODUCT-2	LARANJA
Jarri	jarri	91175000#257557008# 705586008#703503000# 51104005#417924000# 421066005# 1691000175103	FINDING#ENVIRONMENT #OBJECT#OBSERVABLE# QUALIFIER-3	BERDEA
Zaio	izan	257621007#52101004# 304608003#421139008# 260232004	ENVIRONMENT# QUALIFIER-4	HORIA
.	.	-	-	-
Solairuan	solairu	31006001	ORGANISM-5	MOREA

Bi	bi	224857005#19338005# 421832002#1305003# 51440002'	ENVIRONMENT# QUALIFIER-6	HORIA
Aldiz	Aldiz	-	-	-
Asaldata	asaldata	286582007#3646006# 39869006#9251003# 258141001#260379002# 18307000'	FINDING#BODY STRUCTURE#OBJECT# OBSERVABLE# QUALIFIER-7	BERDEA
Da	izan	257621007#52101004# 304608003#421139008# 260232004	ENVIRONMENT# QUALIFIER-8	HORIA
.	.	-	-	-
Tratamendua	tratamendu	119270007#276239002	PROCEDURE-9	URDIN ARGIA
Hasi	hasi	-	-	-
Eta	eta	421417004#421829000# 410513005	QUALIFIER-10	HORIA
Hamar	hamar	3445001	QUALIFIER-11	HORIA
Egunera	egun	421723005#258703001# 15240007	QUALIFIER-12	HORIA
Mesfidantza	mesfidantza	706876006	FINDING-13	BERDEA
Desagertu	desagertu	360304005	OBJECT-14	URDIN ILUNA / MOREA
Zaio	izan	257621007#52101004# 304608003#421139008# 260232004	ENVIRONMENT# QUALIFIER-15	HORIA
,	,	-	-	-
Baita	baita	260521003	QUALIFIER-16	HORIA
Eldarnioak	eldarnio	2776000#71038003#2073000'	DISORDER#FINDING	GORRIXKA
Eta	eta	421417004#421829000# 410513005	QUALIFIER-18	HORIA

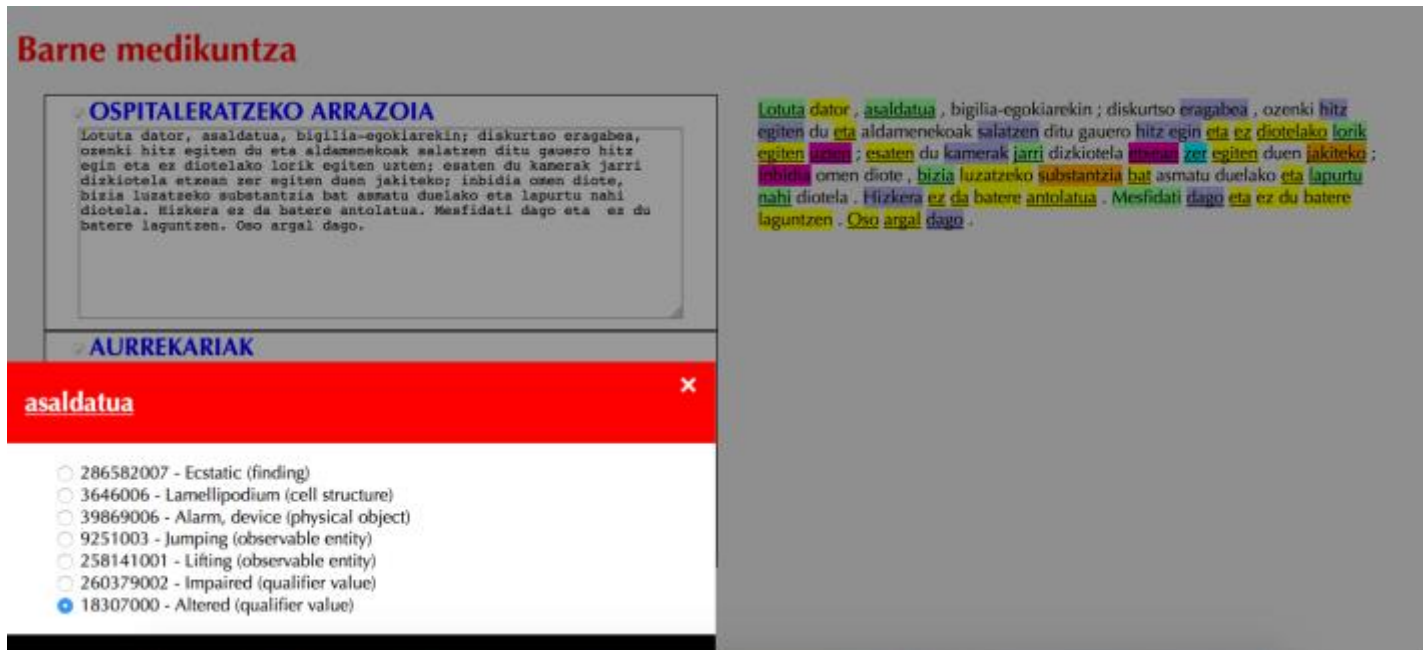
Taula 7. 7: etiketa kolorea - hierarkia baliokidetasuna

Beraz, ikusten den moduan, konkordantzia bat dago taula eta 6.1 irudiko testuan agertzen diren etiketa koloreekin.

**Termino bati dagozkion ID eta FSNak lortzen direla baieztatu:**



Behin interfazea sortua izan dela (6.6 azpiatala) etiketa batean klik egiterakoan termino horren SNOMED ID guztiak eta dagozkien FSN-ak *PopUp* leiho batean agertuko dira (produktuaren helburua betez). Ikusi 7.2 irudia:



Irudia 7. 2: "asaldatuta" terminoari dagozkion SNOMED ID eta FSN-ak

Behin proba guzti hauek eginda, produktua bi ebaluatzailek ebaluatuko dute.

## 7.2 Ebaluazioa

Azpiatal honetan medikuntza arloan jakitunak diren bi lagun osatEUS aplikazioaren ebaluazioa egitea eskatu zaie. Horretarako, terminoak era egokian etiketatuak izan diren ebaluatuko dute. Hau da, termino bat etiketa gorriarekin agertzen bada DISORDER motakoa dela esanahiko du (edo etiketa hau lehenetsia duela besteekin konparatuz). Hau egia bada, termino hau ez da 7.8, 7.9 eta 7.10 tauletan agertuko. Beraz, hurrengo taulan gaizki etiketatuak izan diren terminoak edo analizatzaileak terminotzat detektatu dituenak baino errealitatean ez direnak tauletan agertuko dira. Prozedura hau, bi osasun-txostenekin egin beharko dute bi lagunek sistemaren barruan sartuta daudela (ikusi C eranskinen azkeneko bi osasun-txostenak).

Bi ebaluatzaileen (batek denbora faltarengatik dokumentu bat etiketatu duena eta besteak bi) emaitzak agertuko dira (7.8, 7.9 eta 7.10 taulak). Komuna den dokumenturako (7.8 eta 7.9 taulak) zenbatetan egon diren ados ikusiko da 7.11 taulan. Honi "etiketatzaileen arteko adostasuna" edo *inter annotator agreement* deitzen zaio (7.2.1 azpiatala).

Beraz, hona hemen osasun-txosten bakoitzari dagozkien taulak

### 1. Osasun-txostena- KIRURGIA OROKORRA

Terminoak	Hierarkia egokia	Terminoak	Hierarkia egokia
duela	EZER EZ	dugu	EZER EZ
eta (x11)	EZER EZ	bidezko	QUALIFIER
datorren	EZER EZ	Egin (x2)	EZER EZ
urteko	EZER EZ	diogu (x2)	EZER EZ
azken	EZER EZ	eboluzio	FINDING
dituen	EZER EZ	Izan (x2)	EZER EZ
pazienteak	FINDING	du	EZER EZ
diote	EZER EZ	da eta	EZER EZ
edo (x3)	EZER EZ	dago (x2)	EZER EZ
Heste-soinua	BODYSSTRUCTURE	dagoen	EZER EZ
da (x4)	EZER EZ	markatu	PROCEDURE
tono normala	FINDING	dugun	EZER EZ
simetrikoa	FINDING	herenean	BODYSSTRUCTURE
erritmo	FINDING	ez (x11)	EZER EZ
min	QUALIFIER	metastarik	DISORDER
sailkapenaren	EZER EZ	kokatua	EZER EZ
araberako	EZER EZ		

Taula 7. 8: Kirurgia Orokorreko AnaMed\_eu gaizki detektatutako terminoak 1. ebaluatzailearen arabera.

**EBALUATZAILE 1:** Osasun-txosten honetan hasieran detektatu dituen 187 teminoetatik ebaluatzailearen arabera 126 termino ongi detektatu ditu. Beste hitzetan anaMed\_eu analizatzaileak detektatzearen %67,37 ongi egin du.

Terminoak	Hierarkia egokia	Terminoak	Hierarkia egokia
eta (x11)	EZER EZ	adenopatiarik	DISORDER
egitera	EZER EZ	Goi-ondesteko	BODYSTRUCTURE

ez (x11)	EZER EZ	lerro pektineotik	BODYSTRUCTURE
Zurrungaria	FINDING	ezkerraldean	QUALIFIER
etorri	EZER EZ	kokatua	EZER EZ
dislipemiarik	DISORDER	mesorreктоan	BODYSTRUCTURE
edo (x3)	EZER EZ	operatu	PROCEDURE
diabetesik	DISORDER	dugu	EZER EZ
ez du	EZER EZ	Aurre-erresezioa	PROCEDURE
edaten	EZER EZ	egin (x2)	EZER EZ
behar duzu	EZER EZ	diogu	EZER EZ
da (x4)	EZER EZ	herenean	QUALIFIER
metamizola	PRODUCT	Kolon-ondesteko	BODYSTRUCTURE
enalaprila	PRODUCT	Dago(x2)	EZER EZ
Palexia	PRODUCT	Babes-ileostomia	PROCEDURE
berria	QUALIFIER	behar duzu	EZER EZ
kontrolerako	EZER EZ	du	EZER EZ
Izan (x2)	EZER EZ	sukarrik	FINDING
dituen	EZER EZ	kontrolerako	EZER EZ
eta	EZER EZ	dagoen	EZER EZ
Koloproktologiako	QUALIFIER	dabil	EZER EZ
diote	EZER EZ	Horrenbestez	EZER EZ
C	QUALIFIER	eman	EZER EZ
O2Sat	FINDING	diferentziatua	FINDING
kolorazio	EZER EZ	neoplasikorik	DISORDER
etorri	EZER EZ	inbasio linfobaskularrik	FINDING
dagoen	EZER EZ	Eresezio-ertzak	BODYSTRUCTURE
ultzeraturik	FINDING	neoplasiarik	DISORDER
kendu	PROCEDURE	daude	EZER EZ
Ez dago	EZER EZ	LAPAROSKOPIA	PROCEDURE
ingurgitaziorik	FINDING	Aurre-erresezioa	PROCEDURE
barru	EZER EZ	Metamizola	SUBSTANCE
Uzki-ertzetik	BODYSTRUCTURE	Parasetamola	SUBSTANCE
iododuna	QUALIFIER	metamizolarekin	SUBSTANCE
Da (x4)	EZER EZ	azalpeko injekzio	PROCEDURE
Hatz-punta	BODYSTRUCTURE	bidez	PROCEDURE
Odol-arrastoz	FINDING	garbitu	PROCEDURE
zikindua	FINDING	Kontraindikazio	FINDING
Er ritmo sinusala	OTHER	lehortu	PROCEDURE
Donostia Uni bertsitate Ospitaleko (x2)	OTHER	Koloproktologiako kanpo-kontsultetara	OTHER

Koloproktologiako kanpo-konsultetara	OTHER	Donostia Unibertsitate Ospitaleko	OTHER
---	-------	---	-------

Taula 7. 9: Kirurgia Orokorreko AnaMed\_eu gaizki detektatutako terminoak 2.ebaluatzailearen arabera.

**EBALUATZAILE 2:** Osasun-txosten honetan hasieran detektatu dituen 187 terminoetatik ebaluatzailearen arabera 85 termino ongi detektatu ditu. Beste hitzetan anaMed\_eu analizatzaileak detektatzearen %45,85 ongi egin du. Bigarren ebaluatzailea zorrotzagoa da edo azterketa sakonagoa egin du.

## 2. Osasun-Txostena- ANESTESIOLOGIA ETA BERPIZTEA

Terminoak	Hierarkia egokia	Terminoak	Hierarkia egokia
Eta (x19)	EZER EZ	neurri	OTHER
Egin (x3)	EZER EZ	kontserbatzaileak	OTHER
Aurpegi-masailletako (x2)	BODYSTRUCTURE	balantzatuen	PROCEDURE
Ahoko	BODYSTRUCTURE	zabaltzeko	OTHER
Iekualdatzea	EZER EZ	baiezatu	OTHER
iododunari	QUALIFIER	arnas	FINDING?
Anisakisari	SUBSTANCE	autonomia	FINDING?
Da (x6)	EZER EZ	hemodinamikoki	QUALIFIER
litiasia	DISORDER	dago	OTHER
Kirurgiako	PROCEDURE	Tentsio arteriala	FINDING
matrailaren	BODYSTRUCTURE	Tratamendu basoaktiborik	PROCEDURE
disco-hernia	DISORDER	O2Sat	FINDING
Valsartan	SUBSTANCE	onarekin	QUALIFIER
alopurinola	SUBSTANCE	murriztua	QUALIFIER
Sintrom	SUBSTANCE	haustura	OTHER
Hematologiako	QUALIFIER	Eskuin-birikaren	BODYSTRUCTURE
jarraipen	PROCEDURE	hipofonetiko	FINDING
ankilosia	DISORDER	ahoko bibrazioak ezabatuta	FINDING
hautsi	OTHER	daude	EZER EZ
zaio	EZER EZ	zainpagarria	QUALIFIER
Ankilosaturiko	DISORDER	haztatzean	PROCEDURE
kokatu	OTHER	hipogastroan	BODYSTRUCTURE
ANKILOZIA	DISORDER	ez	EZER EZ

hobi	BODYSTRUCTURE	dauka	EZER EZ
mandibularrean	BODYSTRUCTURE	ARTROPLASTIA	PROCEDURE
egitean	OTHER	ANKILOSIA	DISORDER
tinpanikoa	BODYSTRUCTURE	ARTROPLASTIA	PROCEDURE

Taula 7. 10: Anestesiologia eta Bepizteko AnaMed\_eu gaizki detektatutako terminoak 2.ebaluatzailearen arabera.

**EBALUATZAILE 2:** Osasun-txosten honetan hasieran detektatu dituen 206 teminoetatik ebaluatzaileen arabera 126 termino ongi detektatu ditu. Beste hitzetan anaMed\_eu analizatzaileak detektatzearen %61.17 ongi egin du.

Beraz, emaitzak (C Eranskina) eta 7.8, 7.9 eta 7.10 taulak konparatuta AnaMed\_eu detektatu dituen termino asko ez dira (eta, ez, edo, ...) terminoak. Honen ondorioz, etorkizuneko lan bezala geratzen da termino hauen ezabatzea sct1\_Descriptions1\_eus\_20150731.txt fitxategitik. Gainera, termino batzuk esleitutako hierarkiak gaizki esleituak izan dira edo ebaluatzaileen usteetan termino horiek ez zuten etiketa kolore horrekin agertu behar (ikusi C.2 irudiko erritmo terminoa. Termino honi OsatEUS aplikazioak OTHERS etiketa kolorea ezarri dio eta lehenengo ebaluatzailearen arabera QUALIFIER etiketa esleituko zion). Dena den, emaitza nahiko positiboak lortu dira bigarren osasun-txostenean ehuneko handia (%60a baino gehiago) zuzen etiketatu dituelako. Ez, eta edo da bezalako hitzak soilik kenduta, termino zuzena detektatzearen %70era helduko zen.

### 7.2.1 etiketatzaileen arteko adostuna

Bi etiketatzaileen lehenengo osoasun-txostenen detektatze txarrak emaitzak ikusita, hitz hauetan bat egin dute (7.11 taula).

Hitza	Hitza	Hitza	Hitza
Eta (x11)	Da (x4)	Izan (x2)	Herenean
Dituen	Dugu	Du	Ez (x11)
Diote	Diogu	Dagoen	kokatua
Egin (x1)			

Taula 7. 11: Bi adituen etiketatzaileen arteko adostasuna.

7.11 taulan ikus dezakegun bezala, 27 terminoetan egon dira ados (batzuk errepikatuak) hauen etiketatze okerrean. Esan dezakegu ados egon direneko hitz gehienak aditzak direla eta beraz, hurrengo hau ondorioztatu dezakegu: aditz gehienak ez direla terminoak. Lehenengo ebaluatzaileak ebaluaketa sakon bat egingo baluke, beharbada termino gehiago komunean egongo ziren eta ondorio gehiago lortu genituzkeela.

Hurrengo atalean proiektua gauzatzeko egin den kudeaketa eta jarraipena azalduko da hainbat ondorio azalduz.

# 8

---

## Kudeaketa eta jarraipena

Proiektu guztiaren jarraipena atal honetan egingo da. Gertatu diren arriskuei buruz (8.5 Azpiatala), eskainitako orduak (8.1 azpiatala) etab. azalduko dira.

### 8.1 Burututako lana

Aurretik esan dugun bezala, proiektu hau gauzatzeko 300 ordu erabiltzea pentsatu zen eta ataza bakoitzari eskaini beharreko orduak estimatu

ziren. Ondorengo taulan estimazioak, benetan erabili diren orduak eta bien arteko aldea (desbideraketak) azaltzen dira.

ATAZA	ESTIMATUTAKO ORUDAK	ERABILITAKO ORDUAK	DESBIDERAKETAK	%
Planifikazioa	19:00	13:40	+ 5:20	13.3
Kudeaketa	25:00	22:20	+ 2:40	6.41
Prestakuntza	33:00	25:00	+ 8:00	19.75
Diseinua	13:00	10:10	+ 2:50	6.79
Garapena	90:00	130:30	- 40:30	100
Probak	20:00	19:50	+ 0:10	0.67
Memoria	100:00	121:30	- 21:30	53.1
<b>GUZTIRA</b>	<b>300:00</b>	<b>343:00</b>	<b>- 43:00</b>	<b>106.17</b>

Taula 8. 1: sortutako desbideraketak.

Proiektuan desbideraketa handi xamar bakarria sortu egin da inplementazio atazan. Desbideraketa hau normala da sortzea, produktua sortzerakoan sortu diren arazoengatik (ondoren konponduak izan direnak). Ataza hau kenduta, sortu diren desbideraketa gehienak “positiboak” izan dira, hau da, gehienei espero ziren baino ordu gutxiago eskaini zaizkie.

Beraz, esan dezakegu orduekin estimazio nahiko ona egin dela ataza gehienekin (garapeneko atazan izan ezik, 40:30 orduko desbideraketa sortu baita). Ikusi D eranskina jakiteko zein paketeak ez dira bete finkatutako datan eta beraz, hauek izan dira desbideraketen arrazoi nagusienak.

Ehunekoa honela kalkulatu da: 40,5 zifra desbideraketaren %100 hartu izan da. Beraz, honen arabera besteen ehunekoak kalkulatu izan da.

Behin hau eginda, proiektua gauzatzeko planari buruz hitz egingo da errealitatean bete diren mugarririk edo egutegiari buruz hitz egin.

## 8.2 Proiektua gauzatzeko plana

ATAZA IZENA	ATAZA ZENB
Planifikazioa	1
Kudeaketa	2
Jarraipen kontrola	21
Bilerak	22



<b>Prestakuntza</b>	<b>3</b>
SNOMED CT	31
Kodea ulertu	32
Eskuliburua	33
<b>Diseinua</b>	<b>4</b>
Erabilpen kasuak	41
Klase diagrama	42
Arkitektura	43
<b>Garapena</b>	<b>5</b>
<b>Probak</b>	<b>6</b>
<b>Memoria</b>	<b>7</b>

Taula 8. 2: proiektuaren atazak eta zenbakiak

Ataza hauek planifikazio ataletik (3.atala) lortuak izan dira.

### 8.2.1 Egutegia

Hona hemen proiektuaren egutegia. Egutegi honetan hilabetea, eguna eta astearen zenbakia agertzen da. Honen bidez ikus daiteke zein egunetan zenbatgarren ataza landu den eta zenbat ordu eskaini diren. Egun batean ez bada ezer egin Denbora eta Ataza gelaxkak hutsak egongo dira.

Egutegian ataza bakarra lantzeko pentsatuta dago. Egun batean bat bano gehiago lantzen bada, beste gela batzuk sortuko dira eguna, erabilitako denbora eta ataza azalduz.

Orduaren formatuarentzako hurrengo hau erabiliko da:

	<i>HH:MM</i>
non :	HH = orduak
	MM = minutuak

Testu-kutxa 8. 1: Egutegian orduak idazteko formatua.

Beraz, hurrengo itxura edukiko du egutegiak:

Hilabetea		
Eguna	Denbora	Ataza
Eg 1	1:30	22
Eg 2	Amaitutako ataza	
...	...	...
Eg 31		
<b>HILABETEAN</b>	<b>ORDU KOP.</b>	

Taula 8. 3: ataza eta ordua esleitzeko eskema

Urtarrila		
Eguna	Denbora	Ataza
Urt 1		
Urt 2		
Urt 3		
Urt 4		
Urt 5		
Urt 6		
Urt 7		
Urt 8		
Urt 9		
Urt 10		
Urt 11		
Urt 12		
Urt 13		
Urt 14		
Urt 15		
Urt 16		
Urt 17		
Urt 18		
Urt 19		
Urt 20		
Urt 21		
Urt 22		
Urt 23		
Urt 24		
1	Urt 25	1:30 22
	Urt 26	
	Urt 27	
	Urt 28	
	Urt 29	
	Urt 30	
	Urt 31	
HILABETEAN		1:30

Otsaila		
Eguna	Denbora	Ataza
Ots 1		
Ots 2		
Ots 3		
2	Ots 4	
	Ots 5	2:00 31
	Ots 6	
	Ots 7	
3	Ots 8	3:00 31
	Ots 9	2:00 31
	Ots 10	
	Ots 11	
	Ots 12	2:30 31
	Ots 13	1:30 7
	Ots 14	
4	Ots 15	1:00 32
	Ots 16	4:30 31
	Ots 17	1:30 22
	Ots 18	3:00 7
	Ots 19	
	Ots 20	
	Ots 21	1:30 7
	Ots 22	1:00 7
5	Ots 23	
	Ots 24	1:00 7
	Ots 25	
	Ots 26	
	Ots 27	
Ots 28		
Ots 29		
HILABETEAN		30:00:00

Taula 8. 4: urtarrilean eta otsailean egindako lana

Ots 13	Ots 17	Ots 19
3:00	1:00	1:30
42	32	1

Martxoa			
	Eguna	Denbora	Ataza
7	Mart 1		
	Mart 2		
	Mart 3	2:00	32
	Mart 4		
	Mart 5	1:30	32
	Mart 6	1:00	7
8	Mart 7		
	Mart 8	1:30	22
	Mart 9	2:30	32
	Mart 10	0:40	42
	Mart 11		
	Mart 12		
9	Mart 13		
	Mart 14		
	Mart 15		
	Mart 16	1:30	33
	Mart 17	1:30	7
	Mart 18		
10	Mart 19		
	Mart 20		
	Mart 21		
	Mart 22	0:45	5
	Mart 23		
	Mart 24		
11	Mart 25		
	Mart 26	1:00	7
	Mart 27	4:00	5
	Mart 28	2:30	5
	Mart 29	3:30	1
	Mart 30	0:40	22
Mart 31	1:10	1	
HILABETEAN		29:45:00	

Apirila			
	Eguna	Denbora	Ataza
7	Api 1	3:00	1
	Api 2	5:00	5
	Api 3		
12	Api 4		
	Api 5	1:15	22
	Api 6	3:50	5
	Api 7	2:00	5
	Api 8	2:15	5
	Api 9	2:45	5
	Api 10	2:45	7
14	Api 11	2:00	5
	Api 12		
	Api 13	0:30	5
	Api 14	6:00	5
	Api 15		
	Api 16	3:00	5
	Api 17	4:45	5
15	Api 18	3:15	5
	Api 19	1:45	7
	Api 20		
	Api 21	3:15	1
	Api 22	1:45	5
	Api 23	24:00:00	JSON SORTZEN (zerbitzariak lanean) ETA AKATSAK ZUZENTZEN
	Api 24	24:00:00	
Api 25	24:00:00		
Api 26	24:00:00		
16	Api 27	24:00:00	
	Api 28	20:20	
	Api 29	2:45	7
	Api 30		
HILABETEAN		54:35:00	

Mart 27	Mart 28	Mart 30
1:30	1:15	1:15
7	7	1

Api 5	Api 21	Api 22	Api 24	Api 26
2:45	1:40	0:35	0:40	0:50
5	5	6	7	22

Taula 8. 5: matxoan eta apirilean egindako lana

Maiatza		
Eguna	Denbora	Ataza
17	Mai 1	
	Mai 2	2:00 7
	Mai 3	
	Mai 4	
	Mai 5	1:00 22
	Mai 6	
	Mai 7	0:50 5
	Mai 8	4:30 5
18	Mai 9	4:45 5
	Mai 10	2:30 7
	Mai 11	4:40 5
	Mai 12	3:30 5
	Mai 13	1:40:00 7
	Mai 14	1:40 7
	Mai 15	3:45 7
19	Mai 16	3:15 7
	Mai 17	4:45 5
	Mai 18	3:00 5
	Mai 19	1:00 JSON sortu
	Mai 20	24:00:00
	Mai 21	24:00:00
	Mai 22	14:00
20	Mai 23	2:30 7
	Mai 24	3:00 6
	Mai 25	5:00 7
	Mai 26	5:00 5
	Mai 27	3:00 6
	Mai 28	2:00 6
	Mai 29	8:00 5
	Mai 30	4:30 7
	Mai 31	4:45 7
HILABETEAN		115:20:00

Mai 9	Mai 23
1:00	1:45
7	6
Mai 12	Mai 24
1:15	3:15
22	7
Mai 13	Mai 25
2:15	2:00
5	5
Mai 18	Mai 26
2:45	0:30
43	22
Mai 19	2:30
0:50	7
22	Mai 28
Mai 21	3:30
0:40	7
21	Mai 30
Mai 22	3:15
1:15	5
43	
2:00	
6	
Mai 23	
1:45	
6	
Mai 24	
3:15	
7	
Mai 25	
2:00	

Taula 8. 6: maiatzean egindako lana

Ekaina			
	Eguna	Denbora	Ataza
21	Eka 1	5:00	7
	Eka 2	0:50	22
	Eka 3	3:45	5
	Eka 4	6:00	5
	Eka 5	6:00	7
22	Eka 6	11:30	5
	Eka 7	9:30	5
	Eka 8	4:30	5
	Eka 9	9:00	berritu JSON
	Eka 10	24:00:00	
	Eka 11	12:00	
	Eka 12	2:30	41
	23	Eka 13	5:30
Eka 14		1:00	22
Eka 15			
Eka 16			
Eka 17		2:00	7
Eka 18			
Eka 19			
Eka 20		4:30	7
Eka 21			
24	Eka 22		
	Eka 23		
	Eka 24	AMAIERA	
	Eka 25		
	Eka 26		
	Eka 27		
25	Eka 28		
	Eka 29		
	Eka 30		
HILABETEAN		109:50:00	

Eka1	Eka 12
2:30	1:30
6	33
Eka2	2:15
3:15	6
21	Eka 13
2:00	2:00
7	21
Eka 4	5:30:00
1:00	7
6	Eka 14
Eka 7	4:00
1:00	7
22	
Eka 8	
2:30	
7	
Eka 9	
5:30	
7	
0:45	
22	
Eka 10	
6:00	
7	
Eka 11	
5:30	
7	
2:00	
21	

Taula 8. 7: ekainan egindako lana

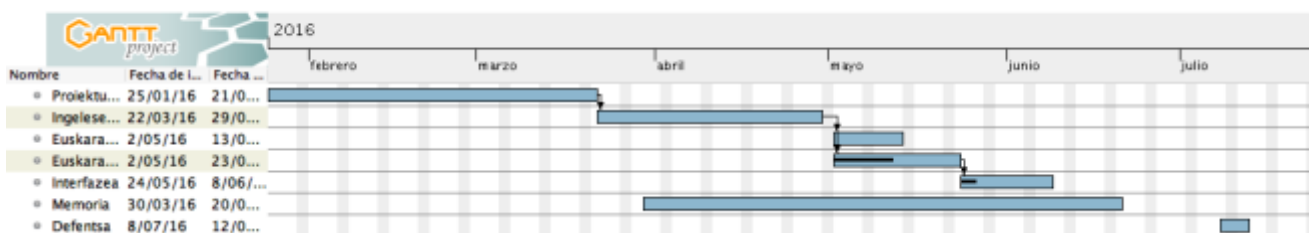
Hemen ikus dezakegu hilabeten zehar zein ataza egin diren eta zenbat denbora eskaini zaizkien. Proiektuaren itxiera **Ekainaren 20**-an izan da, eta espero zen dataren barnean gauzatu da proiektua.

Ikus dezakegu hilabetero ordu kopuruak handitzen joan direla. Bereziki azkeneko bi hilabetetan. GAP-a eskatzen zituen ordu kopurua 300:00 ordukoa bazen 225:10 ordu eskaini zaizkie.

Gainera, ikus dezakegu D eranskinaren bitartez ea bi daten artean finkatu diren lan-paketeak bete diren. Helburu gehienak bete diren arren gutxi batzuk ez dira bete finkatutako denboran. Bete ez diren helburuengatik aurretik ikusi ditugun desbideraketak sortu egin dira (8.1 azpiatala).

### 8.2.2 Gantt diagrama

Orain ikusiko dugun Gantt Diagrama ondoren ikusiko diren mugak jarraituz sortu da.



Irudia 8. 1: Gantt diagrama

**Proiektuaren aurrekariak:** Hemen produktua egin ahal izateko hainbat terminoen barneraketa egin dira (SNOMED CT, AnaMed\_en, TermMed\_eu, ...). Gainera, Med\_en sistemaren klase diagrama eta analisisa egin da.

**Ingeleseko zerbitzaria:** StandFord zerbitzaria hobetu da terminoen informazio lexikoa gordez fitxategi batean.

**Memoria:** Kurtsoan zehar egingo den proiektu guztiaren lana hemen islatu da.

**Euskarazko Zerbitzaria (TermMed\_eu):** TermMed\_en-en euskarazko bertsioa sortu da.

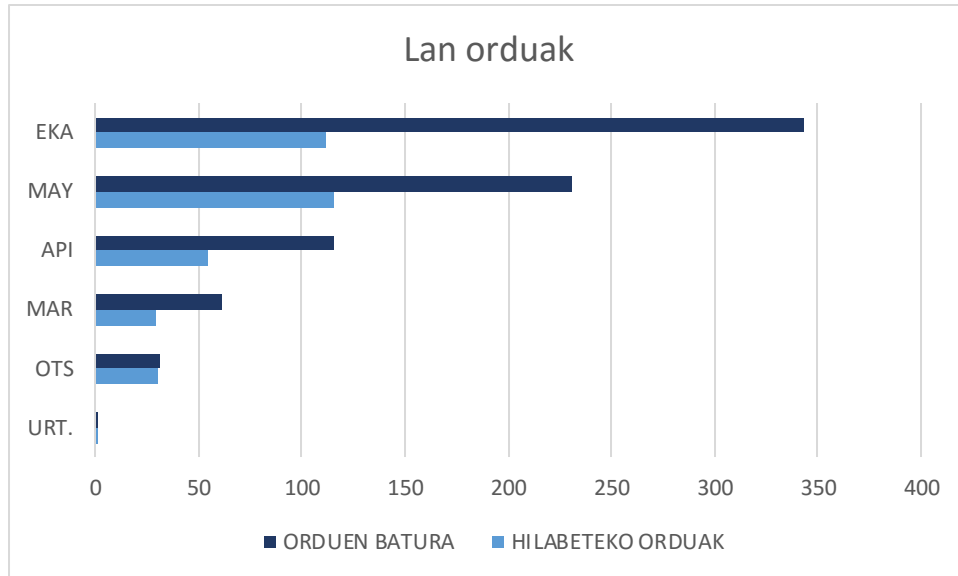
**Euskarazko bezeroa (AnaMed\_eu):** AnaMed\_en-en euskarazko bertsioa sortu da.

**Interfazea:** erabiltzailea egindako lana modu erakargarri batean egin azaltzea.

**Defentsa:** Ikasleak proiektuaren zehar egindako lana aurkeztuko du.

### 8.2.3 Lan-karga

Hona hemen hilabete bakoitzean eskaini zaizkien lan orduak (8.1 grafikoa).



Grafiko 8. 1: Hilabeteko lan orduak

Ikus daitekeen bezala, 8.1 grafikoa bi zutabe ezberdin daude. Urdin argiago den zutabea hilabete bakoitzean sartutako orduak azaltzen dira. Hilabeteak aurrera joan ahala ikusten da nola lan-karga handituz joan den. Hau da, proiektuaren entrega data hurbilduz ahala, lan orduak handituz joan dira espero zen bezala.

Urdin iluna duen zutabearen aldiz, hilabete bakoitzean lan orduen batura agertzen da eta espero zen bezala, zutabe hauek geroz eta azkarrago joan dira “handitzen” aurretik esan dugun bezala, hilabete bakoitzean lan orduak handituz joan direlako.

Beraz, esan dezakegu proiektuaren prestakuntza-fasean (urtarrila, otsaila eta martxoa) ordu gutxiago eskaini zaizkiola inplementazio fasearekin konparatuz (apirila, maiatza eta ekaina).

Hurrengo azpiataletan ikusiko dugu planifikazio atalean zein arrisku, komunikazioa edo kalitate maila bete diren.

## 8.3 Komunikazioa

Komunikazio gehiena posta elektronikoz burutu da. Bilerak (ikus B Eranskina) batez ere mugarrietara iristerakoan egin dira; egindakoa begiratzeko eta nola jarraitu aztertzeko. Orokorrean, komunikatzeko arazorik ez da egon .

Korreo elektronikoz arazo edo dudak txikiak galdetu izan dira. Bileretan aldiz (B eranskina), proiektuaren garapenari buruz hitz egin da egin dena azalduz eta hurrengo helburuak finkatuz.

## 8.4 Kalitatea

Produktuaren kalitatearen inguruan , zerbitzariak eskatutakoa betetzen du, hau da, terminoen analisi egokia egitea SNOMED CTeke terminologia erabiliz. Hainbat proba egin ondoren (7. Atala) lortu behar genukeena lortzen dugu (hitz anitz terminoen eta bakunen analisi egokia) euskaraz eta ingelesez lehenengo saiakeran lortu ez arren.

Bezeroaren kalitatea, berriz, ezin izan da modu kuantitatibo batean neurtu. Probak eskuz burutu dira, behin eta berriz, eskatutakoa ongi betetzen dela jakiteko. Hala ere, bezeroek proba guztiak pasatzen dituela jakiteak konfidantza eman du sistema erabiltzeko orduan. Interfazean lortu dugu espero genituen terminoak identifikatzea C eranskineko testuak idatzi ondoren interfazearen testu-kutxetan (hitz bakunekoak, nahiz hitz anitzekoak).

OsatEUS aplikazioaren erantzuna, aldiz, motela da. Erabiltzaileak 8-9 segundo inguru itxoin beharko du testu-kutxetan idatzi duen testuaren terminoen detektatzea lortzeko. Aplikazioa probatu den hiru nabigatzaileetan arazorik gabe dabil (Firefox, Safari eta Chrome).

Hau idazterako orduan, memoriak 129 orri inguru zituen, 201 eranskinarekin. 100 orritik urrundu egin da, luzapenetan ibili gabe. Beraz, dena ongi azaldu nahi bazen, luzera egokia dela uste da. Memoria tutoreak berrikustez gain, proiektuaren gaia ezagutzen ez duen hirugarren pertsona batek zati batzuk irakurri ditu (1, 2 eta 9 atalak) eta ulergarritasunari buruzko bere iritzia eman du.



## 8.5 Arriskuak

Arrisku bakarria bete izan da proiektuaren zehar. Espero zen bezala, hainbat desbideraketa sortu izan dira, baina azkenean desbideraketei aurre egin izan zaizkie modu arrakastatsu batean. Hona hemen desbideraketa nagusiak:

- **IXA taldean erabiltze gonbidatua izateagatik hainbat baimen ez edukitzea:** honek desbideraketa txikiak sortu ditu. Erabiltzaileak baimen batzuk ez zituen ezin zituen momentuan baliabide batzuk erabili (.txt batzuetako informazioa irakurri, Eustagger\_lite erreminta erabili, ...) eta baimena hauen zain egon behar izan du lanean segi ahal izateko. IXA taldeko "administratzaileei" hainbat baimen eskatzea izan da konponbidea.
- **Eustagger-litek hainbat hitzen lemagizki bueltatu:** analizatzaile honen datu-basean ez dira agertzen hainbat hitz (meningeo, megalia, ...) hitz oso teknikoak direlako medikuntza arloan. Honengatik desbideraketa nahiko handia sortu izan da (4-5 egunekoa). Hitz tekniko hauek eustagger\_lite-k duen datu-basean sartu behar izan dira arazo hau konpondu ahal izateko. Behin konponketa hauek eginda informazio lexikoa biltzen zuten fitxategi (analisiak\_eu.json) berri bat sortua izan da, lema egokiak gordez. Fitxategi hau sortzerakoan 2 eguneko desbideraketa egon da.
- Hasierako RF1 fitxategia (**sct1\_Descriptions1\_eus\_20150731.txt**) oso txikia: .txt hau oso txikia zenez (47000 deskribapen inguru), txostenetan identifikatutako (C eranskinak) hainbat termino asko ez ziren agertzen. Arazo hau konpontzeko gaztelaniako bertsioan euskarazko deskribapenen termino baliokidea bilatu behar izan dugu SNOMED CT *browser*-an. Beraien sctId-a lortu behar izan dugu eskuz eta 89 deskribapen inguru sartuak izan zira eskuz bukaerako emaitza esanguratsua izateko. 6-7 eguneko atzerapena egon da. Ondoren, RF1 fitxategi handiago batekin lan egin da.
- **Django interfazetik AnaMed\_eu ezin deitzea:** hainbat arrazoiengatik (Django *framework*-aren ezagutza oso mugatua izan da arrazoi nagusia) hasiera batean ezin izan da deirik egin. Honen ondorioz, ikasleak hainbat ordu igaro ditu Django-ren inguruan informazioa biltzen bere ingurunea hobeto ulertzeko. Behin hau eginda (eta hainbat proba egin ondoren), AnaMed\_eu erremintarekin konexioa egitea lortu da. 4 eguneko desbideraketa egon da.

Hau da, espero zen bezala inplementazioan lortu da desbideraketarik handiena. Dena den, finkatu den helburua (euskarazko termino identifikatzailea sortzea, alegia) lortu da. Beraz, produktua modu arrakastatsu batean itxi da eta honen ondorioz, proiektu guztia.

Bukatzeko, ez da baliabiderik galdu, IXAko ez zerbitzarrietan ez ikasleak erabiltzen zituen baliabideetan (eramangarri eta USB pertsonalak).

## 8.6 Mugak

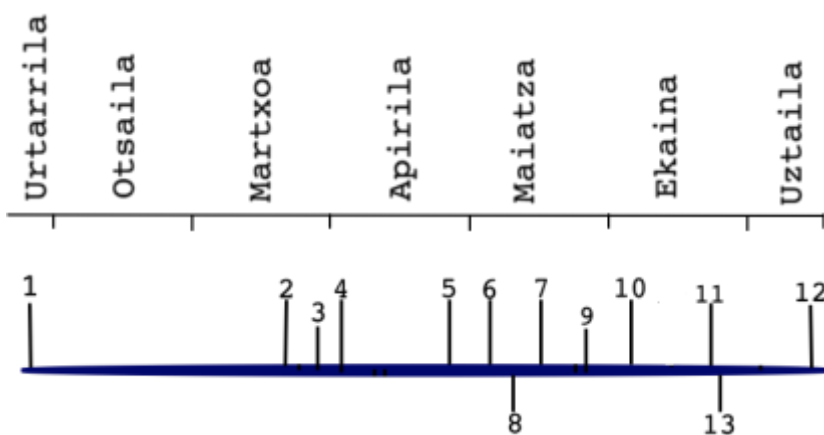
Hona hemen proiektuan betetako benetako mugarriak. Hemengo hauek izan dira errealitatean bete izan diren datak eta ikus dezakegu gutxi batzuk bat egiten dutela hasiera batean estimatuta zeuden mugarriekin.

Planifikazio ataleko eskema berdina segi da, hau da, diagraman hilabeteak agertzen dira eta azpian zenbaki batzuk (8.2 irudia). 8.3 testu-kutxan zenbaki bakoitzari dagokion ataza eta benetako data agertzen dira .

Zenbakia.- Ataza: data estimatua

*Testu-kutxa 8. 2: Mugarriak identifikatzeko egitura.*

Muga berri bat gehitu izan da (13 zenbakia), Interfazea bukatu deiturikoa.



*Irudia 8. 2: Mugarri diagrama*

- 1.- Proiektuaren hasiera: urtarrilak 25
- 2.- Ingeleseko zerbitzaria hobetzen hasi: martxoak 22
- 3.- Memoriaren lehenengo zirriborroa: martxoak 30
- 4.- Planifikazioa: apirilak 2
- 5.- Ingeleseko zerbitzaria hobetzen bukatu: apirilak 29
- 6.- Euskarazko zerbitzaria hasi: maiatzak 5
- 7.- Euskarazko zerbitzaria bukatu: maiatzak 17
- 8.- Euskarazko bezeroa hasi: maiatzak 12
- 9.- Euskarazko bezeroa bukatu: maiatzak 24
- 10.- Interfazea bukatu: ekainak 8
- 11.- Memoriaren azkeneko bertsioa: ekainak 20
- 12.- Defentsa: uztaialak 8-10-11
- 13.- Proiektuaren itxiera: ekainak 23

*Testu-kutxa 8. 3: Mugarriak.*

Ataza asko finkatutako datan ez bukatu arren, batzuk izan duten desbideraketa ez dira handiak izan. Hona hemen ataza bakoitzak eduki duen desbideraketa:

Ataza zenb.	Est. Data (HIL-EGUNA)	Ben. Data (HIL-EGUNA)	Desbideraketa	
			Egunak	%
1	01-25	01-25	0	0
2	03-20	03-22	-2	6.06
3	03-20	03-30	-10	30.3
4	03-28	04-02	- 5	15.15
5	04-08	04-29	- 21	63.63
6	04-09	05-02	- 23	69.69
7	05-20	05-17	+ 3	9.09
8	04-09	05-12	- 33	100
9	05-25	05-24	+ 1	3.03
10	06-15	06-08	+ 7	21.21
11	06-23	06-20	+ 3	9.09
12	07-(08/10/11)	07-(08/10/11)	0	0
13	-	06-23	Ezin jakin	Ezin jakin

*Taula 8. 8: mugarriekin sortu diren desbideraketak.*

Ikusi dugunez, ataza gehienetan desbideraketa txikiak izan dira. Beste hiru atazetan aldiz, espero zena baino desbideraketa handiagoa sortu izan da. Hau da, ez da ondo kalkulatu ataza horiek egiteko behar izango zen egun kopurua bere zailtasunarekiko.

Kasu honetan 33 zifra desbideraketaren %100 hartu izan da. Beraz, honen arabera besteen ehunekoak kalkulatu izan da.

Hau guztia egin ondoren proiektuaren ondorioak eta etorkizuneko lanak azalduko dira produktua hobetzeko asmoarekin.

# 9

---

---

## Ondorioak

Atal honetan GAP-lan honen ondorio nagusiak (9.1 azpiatala) aipatuko ditugu, eta etorkizunerako gelditu zaizkigun atazen (9.2 azpiatala) azalpena emango dugu.

### 9.1 Ondorioak

Lehenik eta behin, Gradu Amaierako Proiektu honetarako ezarri genituen helburuak bete ditugula aipatu beharra dago. Hau da, euskarazko termino identifikatzaile bat sortu da termino motak bereiziz eta emaitza itxaropentsuak jasoz (7. Atala). Honetarako eustagger-lite, SNOMED CT edo NAF bezalako erremintak aztertu behar izan dira hauen erabilera egokia lortzeko. Hau guztia, etorkizuneko lanetan aurrera jarraitzeko bidea irekitzen digu. Aipatu beharra dago espero ziren emaitzak lortu direla termino motak bereiziz (DISORDER, FINDING edo bestelako hierarkiak duten terminoek etiketa kolore bat edo bestea esleituz) eta lehentasunak jarriz termino motei

(DISORDER hierarkia lehenetsun handiago dauka FINDING hierarkiaren aurrean termino bati esleitzerakoan). Gainera, hitz anitzeko terminoen etiketatzea lortu da ere modu egoki batean eta termino anbiguen detektatzea egin da. Terminoak detektatzeko eta horien inguruan informazioa lortzeko interfaze erakargarri bat izatea lortu da, batez ere, *PopUp* leihoaren ondorioz erabili den teknologiarengatik (AJAX, CSS3, HTML5 eta JS). Teknologia honi esker hainbat nabigatzaile desberdinetan OsatEUS aplikazioa erabili daiteke (Safari, Chrome edo Firefox).

Kode gutxi aldatuz (Med\_en sistemak duen kode guztiarekin konparatuz) edozein hizkuntzarako SNOMED CT sistema sor daiteke, euskarazko bertsioaren sorketarekin ikusi dugun bezala. Hori egin ahal izateko, baliabide egokiak eduki behar ditugu (RF1 fitxategiak bezalakoak edo dagozkien analizatzaile linguistikoak) dagozkien hizkuntzan (adibidez, terminoak nahi den hizkuntzan egotea). Sistema ulertzeko, sakon aztertu behar izan da moldaketa egokiak egin ahal izateko.

Proiektua bukatzea ez da erraza izan aurretik sortuta zegoen kodea eta terminologia ulertu eta terminoak barneratu behar zirelako. Gainera, Med\_en hobetzea ere lortu da terminoen bilaketa lehenetsuaren arabera eta datu-basearen sorkuntzarekin sistema bizkorragoa lortu delako (29 orduetatik 2 minutura termMed\_en zerbitzaria martxan jartzea lortu da). Taldean lan egin da (Olatz Perez de Viñasprekin), gerora (enpresen munduan) oso baliagarria izango dena (paraleloki lan egitea, hasieratuta zegoen lana moldatzea,...).

Esan beharra dago, ikasleak planifikatu (3. atala) zituen baino ordu gehiago sartu dituela (43 ordu gehiago, 8. atala) finkatu diren helburuak bete ahal izateko. GAP-a emateko epea hurbildu ahala lan-karga handiagoa izan da, egunero ordu gehiago eskainiz (8.2.1 azpiatala) eta bilerak astero eginez (B eranskina).

Oraindik lan asko egiteke dago euskara – gaztelaniazko txosten elebidunak sortzeko, baino horren sorrera bide egokitik doa egin den lanarengatik.

Beraz, finkatutako helburuak (2. atala) bete dira, nahiz eta proiektu konplexu bat izan.

## 9.2 Etorkizuneko lana

Etorkizunerako lan asko gelditu zaizkigu mahai gainean, nahiz eta jadanik mami askoko lana burututa egon, eta garrantzitsua dena, terminoak euskaraz detektatzea dena. Horren sorkuntza lanaren fase bat besterik ez da izan. Lan guztiaren helburua txostenen itzulpen automatikoa (euskara – gaztelania) sortzea izango litzateke. Horretarako, SNOMED Identifikatzailearen arabera termino baliokideak lortuko genituzke. Hau egin ahal izateko, euskarazko informazioa lexikoa gordetzen duen fitxategia asko handitu beharko genuke gaztelaniaz termino kopuru handiago delako. Behin lan hau eginda, euskara – gaztelania termino baliokideak lortuko genituzke. Gainera, `sct4text` funtzioa hobetu behar da SNOMED identifikatzailetik informazio desberdina lortzeko: FSN, hierarkia edo `PrefTerm`. Beraz, termino batetik hainbat informazioa lortzeko funtzio hau hobeto beharko dugu etorkizun batean (eta informazio hau euskaratu).

Beste etorkizuneko lan bat `eustagger-lite` erreminta hobetzea izango litzateke, horrela lema egokiak lortzeko (gogora dezagun bilaketak lemen arabera egiten ditugula), adjektiboak diren hitz haien izenak lortzeko etab. Horrela, hitz askoren lema egokiak lortuko genituzke (adibidez, paranoide adjektibotik paranoia izena lortu lematzat).

Bukatzeko, termino detektatzea azkarragoa izatea lortu nahi da gerorako (erantzun denbora txikiagoa izatea), honen erantzun denbora 8-9 segundokoa baita eta nahiko motela da.

# 10

---

---

## Bibliografia

### LIBURU EDO ARTIKULUAK:

- Pérez de Viñaspre, O. eta Oronoz, M. (2015).  
Osasun-txosten elebidunak posible ote?
- SNOMED CT Starter Guide. (2014). 2nd ed.
- Perez de Viñaspre Garralda, O. (2013).  
SNOMED CT sare semantikoa euskaratzeko aplikazioa.
- Oronoz Anchordoqui, M. (2008).  
Euskarazko errore sintaktikoak detektatzeko eta zuzentzeko baliabideen garapena: datak, postposizio-lokuzioak eta komunztadura. *Tesi-txostena*, (1), pp.88-89.
- Patrick, J., Wang, Y. and Budd, P. (2016).



An Automated System for Conversion of Clinical Notes into SNOMED Clinical Terminology. In: 1st ed. School of Information Technologies: University of Sydney.

- Perez de Viñaspre Garralda, O. (2016). SNOMED CT sare semantikoa euskaratzeko aplikazioa. *Euskal Herriko Unibertsitatea*.
- Basabe, I. (2014). *Alta Txostenak*. Donostia Unibertsitate Ospitaleak.
- Oronoz, M., Gojenola, K., Pérez, A. and Casillas, A. (2015). Applying Freeling and a Perceptron Sequential Tagger at the Shared Task on Analyzing Clinical Texts. *IXA Taldea (UPV-EHU)*.

#### INTERNETEKO BALIABIDEAK:

- Fdez. Serna, R. (2016). *Medikuntza Graduazio Ekitaldia*. [online] YouTube. Eskuragarri: <https://www.youtube.com/watch?v=vgy12edf9ZU>
- Browser.ihtsdotools.org, (2016). *IHTSDO SNOMED CT Browser*. [online] Eskuragarri: <http://browser.ihtsdotools.org/>
- Stanfordnlp.github.io. (n.d.). *Stanford CoreNLP – a suite of core NLP tools | Stanford CoreNLP*. [online] Eskuragarri: <http://stanfordnlp.github.io/CoreNLP/index.html>
- Wordpress.let.vupr.nl. (2016). *NLP Annotation Format | Background information on NAF*. [online] Eskuragarri: <http://wordpress.let.vupr.nl/naf/>
- GitHub. (2014). *dasmith/stanford-corenlp-python*. [online] Eskuragarri: <https://github.com/dasmith/stanford-corenlp-python>
- Pyspanishdoc.sourceforge.net. (2016). *3.11 pickle -- Serialización de objetos Python*. [online]

- Eskuragarri: <http://pyspanishdoc.sourceforge.net/lib/module-pickle.html>
- Support.google.com. (2016).  
*Work with tab-delimited (tsv) format - Docs editors Help*. [online]  
Eskuragarri: <https://support.google.com/docs/answer/63377?hl=en>
  - Json.org. (2016).  
*JSON*. [online]  
Eskuragarri: <http://json.org>
  - www.tutorialspoint.com. (2016).  
*Python MySQL Database Access*. [online]  
Eskuragarri: [http://www.tutorialspoint.com/python/python\\_database\\_access.htm](http://www.tutorialspoint.com/python/python_database_access.htm)
  - Programmers.stackexchange.com. (2016).  
*Database vs. Pickle (Python)*. [online].  
Eskuragarri: <http://programmers.stackexchange.com/questions/311361/database-vs-pickle-python>
  - Kovshenin, K. (2010).  
*Pickle vs JSON — Which is Faster?*. [online]  
Eskuragarri: <https://kovshenin.com/2010/pickle-vs-json-which-is-faster/>
  - Hellmann, D. (2016).  
*tempfile – Create temporary filesystem resources. - Python Module of the Week*. [online]  
Eskuragarri: <https://pymotw.com/2/tempfile/>
  - Wordpress.let.vupr.nl. (2016).  
*NLP Annotation Format | Background information on NAF*. [online]  
Eskuragarri: <http://wordpress.let.vupr.nl/naf/>
  - Rafalinux.com. (2015).  
*Ejecutar comandos de Linux mediante Python – Rafalinux*. [online]  
Eskuragarri: <http://www.rafaLinux.com/?p=1613>
  - Python?, C. (2016).  
*Converting XML to JSON using Python?*. [online]  
Eskuragarri: <http://stackoverflow.com/questions/191536/converting-xml-to-json-using-python>.

- Docs.python-requests.org. (2016).  
Requests: HTTP for Humans — Requests 2.10.0 documentation.  
Eskuragarri: <http://docs.python-requests.org/en/master/>

# Eranskinak

# *A Eranskina*

---

---

## Erabilitako tresnak

Eranskin honetan erabili diren tresnak guztiak aipatuko dira. Tresna hauek hainbat taldetan sartuak egongo dira hauek zer diren hobeto multzokatzeko. Zertako erabili diren azalduko dira irakurleak garbiago edukitzeko zergatik aipatuak izan diren:

**Nabigatzaleak:** informazioa bilatu, dudak argitu, baliabide berriak aurkitu eta OsatEUS aplikazioa probatu.

- Safari
- Firefox
- Chrome

**Gantt:** Planifikazioan atalean Gantt diagrama egiteko erabilia izan da.

**Paint 2:** Memorian agertzen diren irudi gehienak (ikusitako zerbitzari irudia, edo mugak irudia) *paint* programaren bitartez eginak izan dira.

**Dokumentazioa:**

- Microsoft Word 2016: Memoria idazteko erabili da.

- Microsoft Excel 2016: Ikasleak egunero sartu dituen ataza bakoitzean orduen kontrola egiteko baliagarria eta hainbat grafikoen sorketa.
- Microsoft PowerPoint 2016: Aurkezpena egiterakoan diapositibak sortzeko erabili den baliabidea.

**StarUML** : Analisi atala osatzeko erabilia (Klase diagrama, erabilpen kasuen osaketa)

**GitHub Desktop**: Errepositorio bezala erabilia produktuaren bertsio kontrolak egiteko eta zuzendarientzako eskuragarri egoteko erabilia.

**Terminala**: IXA taldeko zerbitzarietan sartzeko baliagarria. Hemen produktuaren osaketa (kodeketa) egin da.

**Cisco AnyConnect Secure Mobile Client**: Etxetik edo unibertsitatetik ez den beste puntuetatik IXA taldeko zerbitzarietara sartzeko baliagarria.

**Notas**: Bileretan (zuzendari edo zuzendari ordearekin) hainbat ohar hartzeko erabilia.

# *B Eranskina*

---

## **Bilerak**

Atal honetan zuzendari eta laguntzailearekin egin diren bilerak agertuko dira.

Bilera hauen helburuak momentura arte jorratutako gaiak eta hartutako erabakiak idatziko dira proiektuaren kalitatea hobetzeko asmoarekin. Gainera, ikasleak izan dituen desbideraketa nagusiak azalduko dira. Bilerei buruz hitz egiten dugunean fisikoak izango dira, hau da, elkar ikusita, korreoa edo Skype bezalako komunikazio iturriak deuseztatuz.

# Lehenengo Bilera

**Data: 2016ko urtarrilaren 25a**

Hasiera ordua: 12:00

Amaiera Ordua: 13:30

## Gerturatuak:

*Marte Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- *Aurkezpenak*: Maite Oronoz, Olatz Perez de Viñaspre eta Ainhoa Elorz

- *Hasierako egin beharrekoak*:

- SNOMED CT *Started guide* irakurri.
- Klase-diagrama egin `klaseak.py` fitxategiaren bidez.
- VPN, Gantt eta UML Star aplikazioak eramangarri pertsonalean instalatu.
- Python ingurunea ulertu eta instalatu.
- SNOMED CT *browser* begiratu.

- *Proiektuaren helburua*: Osasun-txosten eleanitzak (Euskara- Gaztelania) SNOMED CT erreferentziaz edukita (Euskararen aberastasuna kontuan eduki behar izango du analizatzailean).

## Hartutako erabakiak:

- Aurreko atalean finkatu diren eginbeharrak beteko dira hurrengo bilerarako.

- Proiektuaren azalpen motz bat sortu ikasleak ulertu duena ikusteko.

## Desbideraketak:

Ikasleak portatila konpondu behar izan du, beraz, hiru eguneko atzerapena eduki du eta eginda zeukan dokumentazioa SNOMED CT-ri buruz galdu egin du. VPN berririo instalatu behar izan du. Python ingurunea berriz instalatu behar izan da, beraz, 15 minutu galdu dira instalazioan.



# Bigarren bilera

**Data: 2016ko Otsailaren 17a**

Hasiera ordua: 11:30

Amaiera Ordua: 13:00

**Gerturatuak:**

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

**Jorrotutako gaiak:**

- *Kodea azaldu:* Olatzek ikasleari eginda dagoena azalduko dio.
  - o Klaseak.py sortu diren zalantzak argituak izan dira.
  - o Kodea begiratu eta analizatu.
- *Klase diagrama.*

**Hartutako erabakiak:**

- Inplementazioarekin hasi.
- Klase-diagrama zuzendu behar da (atributu batzuk gehitu, objektuak gehitu, ...).

**Desbideraketak:**

- Klase-diagramaren zuzenketa.

# Hirugarren bilera

**Data: 2016ko Martxoak 8**

Hasiera ordua: 11:00

Amaiera Ordua: 12:40

**Gerturatuak:**

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

**Jorratutako gaiak:**

- *Kodea:*

- o Ikasleari falta zitzaizkien baimenak bereganatu ditu IXA taldeko zerbitzarietan
- o Kode begiratu eta analizatu Olatzekin.

- *Klase diagrama:*

- *Memoria:*

**Hartutako erabakiak:**

- Memorian "Teknologia" atala gehitu erabiliko diren teknologiak azalduz.
- Zein zerbitzaria erabiliko diren azaldu memorian.
- Klase-diagrama zuzendu.
- Kodea sakonago ulertu
- Lemak loturarekin dagoen guztia inplementatu, hau da, aplikazioa probatzerakoan bleed eta bleeding hitzek scid berdina bueltatu behar dute JSONean.

**Desbideraketak:**

- Hasi kodetzen hainbat kontzeptu ez direlako ondo ulertuak izan (kode kantitate asko, azalpen gutxi, ...).
- Enpresa (Deloitte España) batean hainbat azterketa egin behar izanagatik hiru egun ez dira aprobetxatu. Gainera, ikasleak Bilbora bidaiatu behar izan du elkarrizketa pertsonalarengatik.

# Laugarren bilera

**Data: 2016ko martxoak 30**

Hasiera ordua: 11:30

Amaiera ordua: 12:10

## **Gerturatuak:**

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## **Jorratutako gaiak:**

*-Memoria*

*-Kodea*

## **Hartutako erabakiak:**

- *Memoria:* zuzendariak orain arte egindakoari begirada bat botako diote ikasleak egin duen lana ikusteko zein den eta zuzenketak egiteko (hobekuntzak, falta diren zatiak esan). Ikaslea soilik planifikazio atalean ibiliko da.

- *Kodea:* Ikasleak fase berri batekin hasiko da ingeleseko zerbitzaria bukatzeko asmoarekin.

- Ikasleak GitHub kontu bat sortuko du produktua sarean eskuragarri egoteko.

## **Desbideraketak:**

Ez da desbideraketarik sortu.

# Bosgarren bilera

Data: 2016ko apirilaren 5a

Hasiera ordua: 11:30

Amaiera ordua: 12:45

## Gerturatuak:

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

*-Kodea*

## Hartutako erabakiak:

*- Kodea:* Bi erabaki hartu dira:

- klaseak.py fitxategia moldatu behar da ez ain objektuen erabilera malgua izateko.
- Datuak beste baliabide batean kargatu beharko dira .txt-an dauden datuak kargatzerakoan oso motela baita.

## Desbideraketak:

- klaseak.py fitxategian hasieratze, get () eta set () funtzioak sortu klase berri bakoitzerako.

- Datu guztiak kargatzea oso motela da .txt-aren baten bidez (ikusi datu base azpiatala). Datuak gordetzeko baliabide berri bat erabili beharko da kargatzea azkarragoa izateko.

# Seigarren bilera

Data: 2016ko apirilaren 24a

Hasiera ordua: 11:30

Amaiera ordua: 12:20

## Gerturatuak:

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- JSON

- Memoria

## Hartutako erabakiak:

- Kodea:

- JSON: karaktere espezialak direnak (  $CO_2$  bezalakoa, adibidez) normal tratatuko dira ( $CO_2 \rightarrow CO2$  izango da).
- "Q" letra desberdin tratatuko da sortuta zegoen .txt-an gaizki zegoelako eta ez zuen .json fitxategian uzten sartzen honek zuen akatsarengatik.

## Desbideraketak:

- JSON fitxategia hasieratik hasiko sortu izan diren arazoengatik.

# Zazpigarren bilera

**Data:** 2016ko maiatzaren 5a

Hasiera ordua: 11:30

Amaiera ordua: 12:30

## Gerturatuak:

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- *Kodea*: euskarazko zerbitzaria

## Hartutako erabakiak:

- Euskarazko zerbitzaria prestatuko da. Horretarako:

- `eustagger_lite` tresna erabiliko da.
- Ikasleak *Wrapper*-a sortuko du `eustagger_lite` NAF-eratik JSON-era pasatzeko terminoen informazioa.
- Ingeleseko zerbitzaria eredutzat jarraituko da.

## Desbideraketak:

- Ez da desbideraketarik sortu.

# Zortzigarren bilera

**Data: 2016ko maiatzaren 12a**

Hasiera ordua: 11:30

Amaiera ordua: 12:45

## **Gerturatuak:**

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## **Jorratutako gaiak:**

- *Kodea.*

- *Memoria.*

## **Hartutako erabakiak:**

- *Kodea:*

o Euskarazko zerbitzarian aldaketa txiki batzuk egin behar dira, baino berez, amaitutzat dago.

o Euskarazko bezeroarekin hasiko da ikasleak. Horretako AnaMed\_en-eko kodea hartu eta euskararako bertsiorako moldatuko da.

- *Memoria:* Hainbat aldaketak egin behar dira (lehenengo bi kapituluak berregituratu, adibidez)

## **Desbideraketak:**

-Memoria berregituratu.

# Bederatzigarren bilera

**Data: 2016ko maiatzaren 19a**

Hasiera ordua: 9:00

Amaiera ordua: 9:50

## **Gerturatuak:**

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## **Jorratutako gaiak:**

- *Kodea.*
- *Memoria.*

## **Hartutako erabakiak:**

- *Kodea:*
  - o Ikasleak egin duena erakutsi du.
  - o TermMed\_eu-rako .json-aren sorketa egitea erabaki da euskarako deskribapenak erabiliz.
  - o .json-a sortua izan ondoren sisteman hainbat proba egingo dira testu luzeak idatziz (C Eranskineko testuak).
- *Memoria:* Ikasleak zituen hainbat duda argitu dira.

## **Desbideraketak:**

- Ez da desbideraketarik sortu.



# Hamargarren bilera

Data: 2016ko maiatzaren 26a

Hasiera ordua: 11:30

Amaiera ordua: 12:00

## Gerturatuak:

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- *Kodea.*
- *Memoria.*

## Hartutako erabakiak:

- *Kodea:*
  - Interfazearekin hasiko da ikasleak. Horretarako django framework-a erabiliko du. Bukaeran erabiltzaileak idazten duen testutik terminoak detektatuko ditu.
  - Interfazean javascript eta AJAX erabiliko dira terminoak modu ikusgarri batean erakusteko interfazea.
- *Memoria:* Ikasleak zituen hainbat duda argitu dira.

## Desbideraketak:

- *Kodea:*
  - Eustagerekin analisiak egiterakoan ez ditu hainbat hitz errekonozitzen eta beraz, lema gaizki bueltatzen du. Honen zuzenketa egin behar da.
  - Sistemak ez ditu termino luzeenak detektatzen.
  - Badaezpada, ikasleak aurkitu dituen osasun txostenen terminoak berriro begiratuko ditu ea hauen sctId-ak ondo idatzi dituen .txt-an.

# Hamaikagarren bilera

Data: 2016ko ekainaren 2a

Hasiera ordua: 11:30

Amaiera ordua: 12:20

## Gerturatuak:

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- *Kodea.*
- *Memoria.*

## Hartutako erabakiak:

- *Kodea:*
  - o Ikaslea interfazearekin jarraituko du (testua sartzerakoan terminoak detektatzerakoan).
  - o Terminoak kolore bat edukiko du bere hierarkiaren arabera (DISORDER, FINDING, OBJECT, ...). Hierarkia bat baino gehiago dituztenak azpimarratuak agertuko dira.
  - o Aste honetan Eustaggerren egindako zuzenketak (lemak gaizki lortzen zituzten hitzak ondo bueltatu) proba azpiatalean gehituko dira.
- *Memoria:* Planifikazio atala zuzendu egin da.

## Desbideraketak:

- Ikasleak hainbat desbideraketa eduki ditu ezin izan duelako siuc06.si.ehu.es zerbitzaria atzitu Django proiektutik. Arazoa konponduta izan da XQuartz instalatu ondoren MacOS eramangarrian.
- Firefox nabigatzaileira moldatuko da Django aplikazioa, hasiera batean zeuden hainbat *JavaScript* funtzio ez daudelako baliagarri nabigatzaile honetarako (`onFocusOut`, bezalakoa).

# Hamabigarren bilera

**Data: 2016ko ekainaren 9a**

Hasiera ordua: 9:30

Amaiera ordua: 10:30

## **Gerturatuak:**

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## **Jorratutako gaiak:**

*- Kodea*

## **Hartutako erabakiak:**

*- Kodea:*

- Ikaslea falta izan zaizkien hainbat zehaztasun egingo ditu (termino anitzak azpimarratu, anbiguoak diren terminoak azpimarratu).
- Terminoen gainean klik egiterakoan (anbiguoak izanik) hauen sctIdak eta hierakiak agertuko dira.
- Behin hau eginda, FSN-ak lortzen saiatuko gara.

## **Desbideraketak:**

*- Ez da desbideraketarik sortu.*

# Hamahirugarren bilera

Data: 2016ko ekainaren 9a

Hasiera ordua: 11:30

Amaiera ordua: 12:15

## Gerturatuak:

*Maite Oronoz*

*Olatz Perez de Viñaspre*

*Ainhoa Elorz*

## Jorratutako gaiak:

- *Kodea.*

- *Memoria.*

## Hartutako erabakiak:

- *Kodea:*

- Konponketa txiki batzuk egin ondoren produktua amaituta dago.

- *Memoria:*

- Ikasleak inplementazio eta jarraipen eta kontrola talak asko sakonduko ditu.

## Desbideraketak:

- RF1 fitxategi berria behar da termino eguneratuak izateko.

`analisiak_eu.json` fitxategi berria sortuko da fitxategi honen arabera.

# *C Eranskina*

---

## Osasun txosten ereduak

Sei osasun txosten daude (Basabe, 2014). Lehengo lau hauetan (euskarazko bi eta gaztelaniazko beste bi baliokideak direnak), deskribapen anitz agertuko dira hainbat koloreetan: kolore horiak dutenak hasiera batean ondo analizatuak izan direnak izango dira, berde kolorekoak `sct1_Descriptions1_eus_20150731.txt` fitxategian agertzen ez diren deskribapenak ziren (ondoren eskuz fitxategi honetara gehituak izan direnak analisi ona egiteko) eta urdin kolorekoak dutenak `eustagger-lite` analizatzaileak lema gaizki bueltatu duenak izan dira (`eustagger-lite` analizatzailea eraldatu behar izan dena lema egoki bat lortzeko). Gorriz, ikasleak detektatu dituen terminoak baino SNOMED CT-eko *browser*-ean aurkitu ez direnak izan dira. Azpimarratuak izan diren termino guztiak (gorriz izan ezik) 7.1 azpiatalean aurkitzen dira analizaturik.

Falta diren bi txostenekin ebaluatu egingo dira (7.2 azpiatala). Horretarako, medikuntzan adituak diren bi erabiltzaile (`erabiltzaile1` eta `erabiltzaileB` ezizenekin) beharko dira.

ALTAREN TXOSTEN KLINIKOA	
<p>OSPITALERATZEKO ARRAZOA. Agitazioa (asaldura).</p> <p>AURREKARIAK.</p> <p>Orain dela 4 hilabete eskizofrenia paranoidearen tratamendua utzi duen pazientea. Bakarrik bizi da. Bi hilabete daramatza portaera-arazoekin eta azkeneko bi asteetan kontrolik gabe. Ez da bere burua zaintzeko gauza, ez da etxetik ateratzen, eta egun osoan bere gelako pertsianatik begira egoten da. Ez du gaez lorik egiten eta ordutegia guztiz aldatuta dauka. Ez da garbitzen eta gauero bakarrik hitz egiten du ozenki. Aldameneko bizilagunek ez dutela bakean uzten esaten du, eta polizia deitu dio hiru aldiz hori salatzeko. Bizilagunek 112ra deitu dute, eta talde sanitarioa azaldu denean agitatu delako, polizia menderatu egin izan behar du.</p> <p>EGUNGO HISTORIA. Alergia ezagunik ez. Erretzailea: 35 zigarreta eguneko. Ez du hipertentsiorik, ez diabetesik, ez eta dislipemiarik ere. Lau ospitaleratze Akutuentzako Psikiatriako Zerbitzuan eta bi egonaldi ertainekoentzako ospitaleetan. Ohiko tratamendua: Zyprexa Velotab @10 mg (1-0-1); lormetazepama 2 mg, lotarako orduan.</p> <p>AZTERKETA FISIKOA. TA, 131/86 mm Hg; BM, 79 tau/min; AM, 20 arnas/min; O<sub>2</sub>Sat, % 97; T, 38,3 C. Kontziente, orientatua, ondo elikatua eta hidratatua; eupneiko dago. Azal eta mukosetako kolore normala. Burua eta lepoa: ez dauka ingurgitazio jugularrik. Ez dago adenopatiarik. Ez dauka zeinu meningeorik. Biriken auskultazioa: bi biriken aireztapen normala. Ez da entzuten ohiz kanpoko hotsik. Bihotzaren auskultazioa: erritmikoa, foku aortikoan I/VI mailako murmurio sistolikoa. Sabela: biguna eta zanpagarria; minbera eskuineko alboaz haztatzean. Ez dauka peritoneoko narritadurarik; sabel-soinuak normalak dira. Ez da organoen hanturarik nabaritzen, ez eta bestelako masarik ere. Eskuineko giltzurrunaren perkusioa positiboa da. Gorputz-adarrak: edemarik ez. Ez dago zain sakonen tronbosi-zeinunik.</p> <p>AZTERKETA PSIKOPATOLOGIKOA. Lotuta dator, asaldatua, bigilia-egokiarekin; diskurtso eragabea, ozenki hitz egiten du eta aldamenekoak salatzen ditu gauero hitz egin eta ez diotelako lorik egiten uzten; esaten du kamerak jarri dizkiotela etxean zer egiten duen jakiteko; inbidia omen diote, bizia luzatzeko substantzia bat asmatu duelako eta lapurtu nahi diotela. Hizkera ez da batere antolatua. Mesfidati dago eta ez du batere laguntzen. Oso argal dago.</p> <p>PROBA OSAGARRIEN LABURPENA.</p> <p>ANALITIKA: ospitaleratzean desnutrizio orokorra zuen eta altako egunean parametro analitiko guztiak lehengora itzuli dira: glukosa eta ioiak normalak dira. Hemoglobina 13,8; leuzozitoak 5.100; plaketetan eta koagulazioan ez dago neurritz kanpoko baliorik.</p> <p>EBOLUZIOA ETA IRUZKINAK.</p> <p>Ospitaleratzean medikazio antipsikotikoa jarri zaio. Solairuan bi aldiz asaldatu da. Tratamendua hasi eta hamar egunera mesfidantza desagertu zaio, baita eldarnioak eta haluzinazioak ere. Ez du gaixotasunaren kontzientziarik, eta botika hartzea gainbegiratu behar zaio, ezkututzen saiatzen baita, beraz ospitaleratzea luzatzea erabaki da egonaldi ertaineko ospitale batean. San Juan de Dios Ospitale Psikiatrikora lekualdatu da lau hilabeterako.</p> <p>DIAGNOSTIKO NAGUSIA.</p> <p><b>Eskizofrenia paranoidea</b></p> <p>TRATAMENDUA.</p> <ul style="list-style-type: none"> <li>Zyprexa Velotab, 10 mg pilula bat gosarian eta beste bat afarian.</li> <li>Lormetazepama, 2 mg, pilula bat lotarako orduan.</li> </ul> <p>BESTE GOMENDIO BATZUK.</p> <ul style="list-style-type: none"> <li>Ondo orekatu ondoren, epe luzeko tratamendu antipsikotikoa.</li> </ul>	
161	Sinadura-data1/1 orrialdea

Digestivo .

**Motivo de consulta** . Varón de 55 años que presenta astenia .

**Antecedentes personales** . No alergias conocidas . No tratamiento habitual .

**Antecedentes personales** . Hace 3 meses siente astenia . Desde hace unos días le parece hacer deposiciones negruzcas . Después de consultar con el médico de cabecera , le hacen análisis y se confirma la anemia . No siente dolor abdominal ni tampoco hay alteración en el ritmo deposicional .

**Examen físico** . TA-130/80 T-36,7 FC-70 . Consciente , orientado , bien hidratado y perfundido , eupneico , coloración pálida de piel y mucosas . Cabeza y cuello : sin alteraciones patológicas . AP-murmullo vesicular normal . Abdomen blando y depresible , no doloroso a la palpación , no masas ni megalias , ruidos abdominales normales . Extremidades : no edemas , pulsos periféricos adecuados . Tacto rectal : no se palpa masa alguna .

**Resumen de pruebas complementarias** .

Analítica ( en el ambulatorio ) . Hemoglobina-9,6 ; volumen corpuscular medio ( VCM )-99 ; hierro-15 .

EKG-Ritmo sinusal normal .

RXtórax normal .

Gastroscofia . Después de poner anestesia local , se introduce el gastroscopio por boca . Esófago de calibre y mucosa normales . Estómago de extensibilidad y contractibilidad normales , no habiendo restos de sangre ni rastros de otro tipo de lesión . Píloro normal . Bulbo duodenal y segunda porción del duodeno normales .

Colonoscopia . Examen de ano normal . Tacto rectal normal . Recto normal . Colon descendente normal . Colon transverso normal . Colon ascendente normal . Ciego : neoplasia ulcerada y se toma biopsia .

Tomografía abdominal . Hígado , bazo , páncreas , vesícula biliar , vías biliares , suprarrenales y riñones normales . Se observa un engrosamiento de la pared del ciego , de aspecto neoplásico . No hay adenopatía .

Anatomía Patológica . Adenocarcinoma .

**Evolución y comentarios** . Nada más hospitalizarle se la pone a dieta absoluta . No más melenas , y poco a poco se le repone la dieta oral . Se le da hierro intravenoso y mejora la astenia . Una vez diagnosticado de cáncer de ciego se la da de alta , y se le deriva al Servicio de Cirugía General a la espera del tratamiento quirúrgico .

**Diagnóstico principal** . Neoplasia de ciego . Anemia ferropénica .

**Tratamiento** . Se le avisará por teléfono para consulta en Cirugía General .

**Otras recomendaciones** . Evitar estreñimiento , tomando abundante agua y fibra .

**DIGESTIO-APARATUA**

OSPITALERATZEKO ARRAZOA. **Anemia** duen 55 urteko gizonezkoa.

AURREKARIAK. Ez du **alergiarik**. Ez du **ohiko tratamendurik**.

EGUNGO HISTORIA. Duela 3 hilabetetik **ahultasuna** sentitzen du. Egun batzuetatik hona **eginkariak beltzak** direla iruditu zaio. Oheburuko medikuari kontsultatu ostean, **odol-analisia** egin diote eta **anemia** duela ikusi da. Ez du **sabeleko minik** sentitzen, eta ez du izan eginkarien erritmo-aldaketarik.

AZTERKETA FISIKOA. Tentsio arteriala (TA), 130/80 mm Hg; **temperatura**, 36,7 C; bihotz-maiztasuna (BM), 70 tau/min. **Kontziente** eta **orientatua** dago. **Hidratazioa** eta perfusioa egokiak dira. **Eupneikoa**. **Larruazal** eta mukosen zurbiltasuna nabaritzen zaio.

**Burua eta lepoa**. Ez dago aurkikuntza patologikorik.

**Biriken auskultazioa**. Biriketako murmurio normala.

**Sabela**. **Biguna** eta **zanpagarria**. **Haztapenean ez du minik** sentitzen, eta ez dago masa edo megaliarik. **Heste-soinu** normalak.

**Gorputz-adarrak**. Ez du **edemarik**. **Pultsu periferiko** normalak.

**Ondeste-ukipena**. Ez da masarik ukitzen. Eskularruaren hatza melenaz zikindu da.

PROBA OSAGARRIEN LABURPENEA

**ANALITIKA** (ambulatorioan). **Hb**, 9,6 g/dl; **batez besteko bolumen korpuskularra** (BBK), 69; **burdina**, 15 ug/dl.

**ELEKTROKARDIOGRAMA (EKG)**. **Erritmo sinusal normala**.

**BULARREKO ERRADIOGRAFIA**. Normala.

**GASTROSKOPIA**. **Anestesia topikoa** jarri ostean, **gastroskopia** ahotik barrena sartuta: **hestegorriak** kalibre eta **mukosa** normala du; urdailaren zabalgarritasuna eta uzkurtasuna normalak dira; ez da odol-aztarnarik ikusten, ez eta **lesiorik** ere; **piloro** normala. **Duodeno-erraboila** eta duodenoaren bigarren zatia normalak dira.

**KOLONOSKOPIA**. **Uzkiaren ikuskapena**: normala; **ondeste-ukipena**: normala; **ondeste normala**; **sigma**: dibertikuluak; **beheranzko kolon** normala; **zeharkako kolon** normala; **goranzko kolon** normala; **itsua**: **neoplasia** ultzeratua. **Biopsia** egin zaio.

**ABDOMENEO OTA**. **Gibela**, **barea**, **pankrea**, **behazuna**, **behazunbidea**, giltzurrungaineko guruina eta **giltzurrunak** normalak. Itsuaren paretaren gizontze bat ikusten da, **neoplasia** itxurakoa. Ez da adenopatiarik ikusten.

**ANATOMIA PATOLOGIKOA**. **Adenokartzinoma**.

EBOLUZIOA ETA IRUZKINAK. Pazientea ospitalera iritsi bezain pronto, erabateko dieta ezarri zaio. Ez da melena gehiagorik azaldu eta aho bidezko dieta pixkanaka hasteko agindu zaio. Bena barneko **burdina** injektatuz tratatu da eta ahultasunak hobera egin du. **Itsuko neoplasia** diagnostikatu eta Kirurgia Orokorreko Zerbitzuarekin harremanetan jarrita, pazienteari alta eman diogu, tratamendu kirurgikoaren zain.

DIAGNOSTIKO NAGUSIA(K)

**Heste itsuko neoplasia**. **Anemia ferropenikoa**.

TRATAMENDUA

- . **Dieta normala**
- . **Telefonoz abisatuko diote Kirurgia Orokorreko kontsultara joan dadin.**

BESTE GOMENDIO BATZUK

**Idorria** ekidin, **ur** eta **zuntz** ugari hartuz.



**Motivo de hospitalización** : agitación .

**Antecedentes** . Paciente que hace 4 meses ha abandonado el tratamiento de esquizofrenia paranoide . Vive sólo . Lleva 2 meses con problemas de conducta y las 2 últimas semanas está sin control alguno . No es capaz de cuidarse a sí mismo , no sale de casa y pasa el día mirando a través de la persiana de su habitación . No duerme por la noche y su horario está totalmente trastornado . No se lava y por la noche habla sólo a voz en grito . Dice que sus convecinos no le dejan en paz y que ha llamado 3 veces a la policía para denunciarlo . Los vecinos han llamado al 112 y cuando el equipo sanitario ha acudido al lugar ha tenido que intervenir la policía para reducir al paciente a causa de su agitación.

**Historia actual** . No antecedentes de alergia conocidos . Fumador de 35 cigarrillos/día . No hipertensión , ni diabetes ni dislipemia . Cuatro hospitalizaciones en el Servicio de Psiquiatría de Urgencias y dos en postagudos o de media estancia . Tratamiento habitual : Zyprexa Velotab 10 mg. ( 1-0-1 ) , lormetazepam 2 mg./ noche .

**Exploración física** . TA 131/86 . FC 79 pulsaciones/minuto . FR 20 respiraciones/minuto . Saturación O2 97% . T 38,3 grados centígrados . Consciente , orientado , bien alimentado e hidratado , eupneico . Coloración normal de piel y mucosas . Cabeza y cuello . No ingurgitación yugular , no adenopatías , no signos meníngeos . *Auscultación pulmonar* . Oxigenación normal de ambos pulmones , no se oyen ruidos anormales . *Auscultación cardíaca* . Rítmico con murmullo sistólico grado I-IV en foco aórtico . *Abdomen* . Blando y depresible , doloroso a la palpación en flanco derecho . No signos de irritación peritoneal , los ruidos abdominales son normales . No se observan signos de inflamación orgánica ni de masa tumotal . Percusión positiva de riñón derecho . *Extremidades* . No edemas ni signos de trombosis .

**Exploración psicopatológica** .Viene atado , agitado , estado vigil , discurso disforme e irregular , habla a gritos , ha denunciado a los vecinos porque a las noches hablan y no le dejan dormir . Dice que le han puesto cámaras en su casa para saber lo que hace , que le tienen envidia porque ha descubierto una sustancia para alargar la vida y que se la quieren robar . El habla está desorganizado . Desconfiado y suspicaz , no colabora . Muy delgado .

**Resumen de pruebas complementarias** .Si en la analítica hay desnutrición general al ingreso , todos los parámetros analíticos están normalizados al alta . Glucosa y iones normales . Hemoglobina 13,8 , leucocitos 5100 , plaquetas y demás factores de la coagulación normales .

**Evolución y comentarios** . Al ingreso se le ha puesto medicación antipsicótica . Dos veces ha tenido crisis de agitación en planta . A los 10 días de iniciado el tratamiento , desaparece la desconfianza así como los delirios y alucinaciones . No hay anomalías de conciencia . Se debe supervisar la toma de medicación puesto que trata de ocultarla . Por todo ello se decide la hospitalización en un hospital de media estancia o de postagudos durante 4 meses trasladándolo al Hospital Psiquiátrico de San Juan de Dios .

**KIRURGIA OROKORRA**

**OSPITALERATZEKO ARRAZOA.** Ondesteko neoplasia duela eta, ebakuntza kirurgiko programatua egitera datorren 62 urteko **gizonezkoa/emakumea**.

**AURREKARI PERTSONALAK.** Alergiarik ez. Zurrungaria. Uzki-**arteka**. Loaren apnea buxatzailearen sindromea. Ez hipertentsiorik, dislipemiarik edo diabetesik. Ez du alkoholik edaten. Ez da erretzailea. Ez du aurrekari kirurgikorik. **Ohiko tratamendua:** enalaprila, metamizola, Palexia (opioide berria).

**EGUNGO HISTORIA.** Azken hilabeteetan, eginkarietan odola eta anemia ferropenikoa izan dituen pazientea, eta Koloproktologiako Unitateko kontsultan ondesteko tumorea diagnostikatu diote

**EBAKUNTZA AURREKO AZTERKETA.** Tentsio arteriala, 140/80 mm Hg; temperatura, 36,8 C; bihotz-maiztasuna, 75 tau/min; arnas maiztasuna, 17 arnas/min; O<sub>2</sub>Sat basala, % 99. Kontziente eta orientatua. Hidratazio eta perfusio egokiak, **eupneiko**. Mukosa eta larruzaletako kolorazio normala. **Burua eta lepoa.** Ez dago jugularren ingurgitaziorik, ez adenopatia zerbikalik, lepauztai gainekorik edo galtzarbekorik ere. **Toraxa.** **Bihotzaren auskultazioa:** erritmikoa, murmuriorik gabe; **birika-auskultazioa:** murmurio normala. **Sabela.** Biguna, erraz beheratzen dena, haztatzean ez-mingarria, ez peritoneo-narritadurarik, ezta masa edo megaliarik ere; heste-soinu normalak. **Ondeste-haztapena.** Ez da masarik haztatzen, tonu normala, simetrikoa; hatz-punta odol-arrastoz zikindua.

**PROBA OSAGARRIAK. ELEKTROKARDIOGRAMA.** Erritmo sinusala, 75 tau/min. **TORAXEKO ERRADIOGRAFIA** normala. **ANESTESIA AURREKO EBALUAZIOA.** Kontraindikazio anestesikorik ez, ASaren sailkapenaren araberako arrisku anestesikoa: II. **KOLONOSKOPIA.** Uzki-ertzetik 10 cm-ra ultzeraturik dagoen neoformazio bat dago, **zeinari** biopsia egin, eta Txinako tintaz markatu dugun. **Anatomia Patologikoa:** adenokartzinoma. **SABEL ETA PELBISEKO OTA.** Ondesteko neoplasia erdiko herenean. Ez da gibleko metastasirik ikusten, ez tamaina patologikoko adenopatiarik ere. **PELBISEKO ERRESONANTZIA MAGNETIKO-NUKLEARRA.** Goi-ondesteko neoplasia, lerro pektineotik 10 cm-ra, ezkerraldean kokatua. Adenopatia bakar bat **mesorreктоan**, T2 estadio erradiologikoan.

**EBAKUNTZA KIRURGIKOA. 2012ko azaroaren 27an,** anestesia orokorrarekin operatu dugu: laparoskopia bidezko ondestearen aurre-erresekzioa egin diogu, **eta** kolon-ondesteko anastomosi mekanikoa **eta** babes-ileostomia.

**EBOLUZIOA.** Eboluzio ona izan du, sukarrrik gabe, ahotiko dieta berrezarri da eta ileostomia ondo dabil. Horrenbestez, ospitaleko alta eman diogu.

**ANATOMIA PATOLOGIKOA.** Adenokartzinoma ondo diferentziatua (G1), 4x2,5 cm-koa, **afektazio** neoplasikorik gabeko 15 adenopatia, ez dago inbasio linfobaskularrik, erresekzio-ertzak neoplasiarik gabe daude; **pT2N0M0**.

**DIAGNOSTIKOA. ONDESTEKO MINBIZIA (T2N0M0); LAPAROSKOPIA BIDEZKO ONDESTEAREN AURRE-ERRESEKZIOA. ILEOSTOMIA.**

**TRATAMENDUA.** Metamizola 575 mg, 2 kapsula **8 orduan behin**. Parasetamola 1g, **8 orduan behin**, metamizolarekin tartekatuz. Enoxaparina 40 mg eguneko, azalpeko injekzio bidez, hilabetez.

**GOMENDIOAK.** Dutzatzean, garbitu zauriak **ur eta xaboiz**; zauria ondo lehortu eta antiseptiko bat -pobidona iododuna- eman. Zaurietako puntuak 14 egun barru kendu osasun-zentroan. Donostia Unibertsitate Ospitaleko Koloproktologiako kanpo-kontsultetara etorri behar duzu kontrolerako, 2014ko urtarrilaren 8an, 13:00etan.

Sinadura-  
data 1/1  
orrialdea

Hona hemen Kirurgia Orokorreko Osasun-txostenaren terminoen etiketatze emaitzak:

<p>✓ <b>OSPITALERATZEKO ARRAZOAIA</b></p> <p>Ondesteko neoplasia duela eta, ebakuntza kirurgiko programatua egitera datorren 62 urteko gizonzkoa/emakumea.</p>	<p>Ondesteko neoplasia duela eta, ebakuntza kirurgiko programatua egitera datorren 62 urteko gizonzkoa / emakumea .</p>
<p>✓ <b>AURREKARIAK</b></p> <p>Alergiarik ez. Zurrungaria. Uzki-arterka. Loaren apnea buxatzailearen sindromea. Ez hipertentsiorik, dislipemiarik edo diabetesik. Ez da alkoholik edaten. Ez da erretzaila. Ez da aurrekari kirurgikorik. Ohiko tratamendua: enalapril, metamisola, Palexia (opioide berria).</p>	<p>Alergiarik ez. Zurrungaria. Uzki-arterka. Loaren apnea buxatzailearen sindromea. Ez hipertentsiorik, dislipemiarik edo diabetesik. Ez du alkoholik edaten. Ez da erretzaila. Ez du aurrekari kirurgikorik. Ohiko tratamendua: enalapril, metamisola, Palexia (opioide berria).</p>
<p>✓ <b>EGUNGO HISTORIA</b></p> <p>Azken hilabeteetan, eginkarietan odola eta anemia ferropenikoa izan dituen paziente, eta Koloproktologiako Unitateko kontsultan ondesteko tumorea diagnostikatu diote</p>	<p>Azken hilabeteetan, eginkarietan odola eta anemia ferropenikoa izan dituen paziente, eta Koloproktologiako Unitateko kontsultan ondesteko tumorea diagnostikatu diote</p>

Eranskina C. 2: Kirurgia Orokorreko osasun-txostenaren lehenengo zatia.

<p>✓ <b>EBAKUNTZA AURREKO AZTERKETA</b></p> <p>Tentsio arteriala, 140/80 mm Hg; temperatura, 36,8 C; bihotz-maiztasuna, 75 tau/min; arnas maiztasuna, 17 arnas/min; O2Sat basala, % 99. Kontziente eta orientatua. Hidratazio eta perfusio egokiak, eupneiko. Mukosa eta larruazaleko koloreak normala.</p>	<p>Tentsio arteriala, 140/80 mm Hg; temperatura, 36,8 C; bihotz-maiztasuna, 75 tau / min; arnas maiztasuna, 17 arnas / min; O2Sat basala, % 99. Kontziente eta orientatua. Hidratazio eta perfusio egokiak, eupneiko. Mukosa eta larruazaleko koloreak normala.</p>
<p>✓ <b>Burua eta lepoa.</b></p> <p>Ez dago jugularren ingurgitatziorik, ez adenopatia zerbikalik, lepauzta gainerik edo galtzarbekorik ere.</p>	<p>Ez dago jugularren ingurgitatziorik, ez adenopatia zerbikalik, lepauzta gainerik edo galtzarbekorik ere.</p>
<p>✓ <b>Toraxa.</b></p> <p>Bihotzaren auskultazioa: erritmikoa, murmuriorik gabe; biriken auskultazioa: murmurio normala</p>	<p>Bihotzaren auskultazioa: erritmikoa, murmuriorik gabe; biriken auskultazioa: murmurio normala</p>
<p>✓ <b>Sabela.</b></p> <p>Biguna, erraz beheztatzen dena, haztatzean ez-mingarria, ez peritoneo-narritadurarik, ezta masa edo megaliarik ere; heste-soinu normalak.</p>	<p>Biguna, erraz beheztatzen dena, haztatzean ez-mingarria, ez peritoneo-narritadurarik, ezta masa edo megaliarik ere; heste-soinu normalak.</p>
<p>✓ <b>Ondeste-haztapena.</b></p> <p>Ez da masarik haztatzen, tonu normala, simetrikoa; hatz-punta odol-arrastoz zikindua.</p>	<p>Ez da masarik haztatzen, tonu normala, simetrikoa; hatz-punta odol-arrastoz zikindua.</p>

Eranskina C. 1: Kirurgia Orokorreko osasun-txostenaren bigarren zatia.

<p>✓ <b>PROBA OSAGARRIEN LABURPENA</b></p>	
<p>✓ <b>Elektrokardiograma (EKG).</b></p> <p>Erritmo sinusala, 75 tau/min</p>	<p>Erritmo sinusala, 75 tau / min</p>
<p>✓ <b>Toraxeko erradiografia.</b></p> <p>Normala</p>	<p>Normala</p>
<p>✓ <b>Anestesia aurreko ebaluazioa.</b></p> <p>Kontraindikazio anestesikorik ez, ASAren sailkapenaren arabera arrisku anestesiakoa: II.</p>	<p>Kontraindikazio anestesikorik ez, ASAren sailkapenaren arabera arrisku anestesiakoa: II.</p>
<p>✓ <b>Kolonoskopia.</b></p> <p>Uzki-ertzetik 10 cm-ra ultzeraturik dagoen neoformazio bat dago, zeinari biopsia egin, eta Txinako tintaz markatu dugun. Anatomia Patologikoa: adenokartzinoma.</p>	<p>Uzki-ertzetik 10 cm-ra ultzeraturik dagoen neoformazio bat dago, zeinari biopsia egin, eta Txinako tintaz markatu dugun. Anatomia Patologikoa: adenokartzinoma.</p>
<p>✓ <b>Sabel eta pelbiseko OTA.</b></p> <p>Ondesteko neoplasia erdiko herenean. Ez da gibeledoko metastasirik ikusten, ez tamaina patologikoko adenopatiarik ere.</p>	<p>Ondesteko neoplasia erdiko herenean. Ez da gibeledoko metastasirik ikusten, ez tamaina patologikoko adenopatiarik ere.</p>
<p>✓ <b>Pelbiseko erresonantzia magnetikoa.</b></p> <p>Goi-ondesteko neoplasia, lerro pektineotik 10 cm-ra, ezkerrean kokatua. Adenopatia bakar bat mesorrektoan, T2 estadio erradiologikoan.</p>	<p>Goi-ondesteko neoplasia, lerro pektineotik 10 cm-ra, ezkerrean kokatua. Adenopatia bakar bat mesorrektoan, T2 estadio erradiologikoan.</p>

Eranskina C. 3: Kirurgia Orokorreko osasun-txostenaren hirugarren zatia.

<p>✓ <b>EBAKUNTZA KIRURGIKOA</b></p> <p>2012ko azaroaren 27an, anestesia orokorrarekin operatu dugu: laparoskopia bidezko ondestearen aurre-erresekzioa egin diogu, eta kolon-ondesteko anastomosi mekanikoa eta babes-ileostomia.</p>	<p>2012ko azaroaren 27an, anestesia orokorrarekin operatu dugu: laparoskopia bidezko ondestearen aurre-erresekzioa egin diogu, eta kolon-ondesteko anastomosi mekanikoa eta babes-ileostomia.</p>
<p>✓ <b>EBOLUZIOA ETA IRUZKINAK</b></p> <p>Eboluzio ona izan du, sukarririk gabe, ahotiko dieta berrezarri da eta ileostomia ondo dabil. Horrenbestez, ospitaleko alta eman diegu.</p>	<p>Eboluzio ona izan du, sukarririk gabe, ahotiko dieta berrezarri da eta ileostomia ondo dabil. Horrenbestez, ospitaleko alta eman diegu.</p>
<p>✓ <b>ANATOMIA PATOLOGIKOA</b></p> <p>Adenokartzinoma ondo diferentziatua (G1), 4x2,5 cm-koa, afektazio neoplasiakorik gabeko 15 adenopatia, ez dago inbasio linfobaskularrik, erresekzio-ertzak neoplasiarik gabe daude; pT2N0M0.</p>	<p>Adenokartzinoma ondo diferentziatua ( G1 ), 4x2, 5 cm-koa, afektazio neoplasiakorik gabeko 15 adenopatia, ez dago inbasio linfobaskularrik, erresekzio-ertzak neoplasiarik gabe daude; pT 2 N0M0.</p>
<p>✓ <b>DIAGNOSTIKO NAGUSIA(K)</b></p> <p>ONDESTERO MINBIZIA (T2N0M0); LAPAROSKOPIA BIDEZKO ONDESTEAREN AURRE-ERRESEKZIOA. ILEOSTOMIA.</p>	<p>ONDESTERO MINBIZIA ( T2N0M0 ); LAPAROSKOPIA BIDEZKO ONDESTEAREN AURRE-ERRESEKZIOA. ILEOSTOMIA.</p>
<p>✓ <b>TRATAMENDUA</b></p> <p>Metamizola 575 mg, 2 kapsula 8 orduan behin. Parasetamola 1g, 8 orduan behin, metamizolarekin tartekatuz. Enoxaparina 40 mg eguneko, azalpeko injezio bidez, hilabetez.</p>	<p>Metamizola 575 mg, 2 kapsula 8 orduan behin. Parasetamola 1g, 8 orduan behin, metamizolarekin tartekatuz. Enoxaparina 40 mg eguneko, azalpeko injezio bidez, hilabetez.</p>
<p>✓ <b>GOMENDIOAK</b></p> <p>Dutxatzean, garbitu zauriak ur eta xaboiz; zauria ondo lehortu eta antiseptiko bat -povidona iododuna- eman. Zaurietako puntuak 14 egun berriz kendu osasun-zentroan. Donostia Unibertsitate Ospitaleko Koloproktologiako kanpo-kontsultetara etorri behar duzu kontrolerako, 2014ko urtarrilaren 8an, 13:00etan.</p>	<p>Dutxatzean, garbitu zauriak ur eta xaboiz; zauria ondo lehortu eta antiseptiko bat -povidona iododuna- eman. Zaurietako puntuak 14 egun berriz kendu osasun-zentroan. Donostia Unibertsitate Ospitaleko Koloproktologiako kanpo-kontsultetara etorri behar duzu kontrolerako, 2014ko urtarrilaren 8an, 13:00etan.</p>

Eranskina C. 4: Kirurgia Orokorreko osasun-txostenaren azkeneko zatia.

## ANESTESIOLOGIA ETA BERPIZTEA

**OSPITALERATZEKO/KONTSULTARAKO ARRAZOIA.** Masailezurraren eta tenporalaren arteko artroplastia egin ondoren, lekualdaturiko 67 urteko gizonezkoa.

**ALTAREN ARRAZOIA.** Aurpegi-masailetakoa eta ahoko Kirurgiako Zerbitzura lekualdatzea.

**AURREKARIAK.** Alergia die penizilinari, pobidona iododunari eta *Anisakisari*; erretzaile ohia da, hipertentsio arteriala du eta giltzurruneko gutxiegitasun arina; litotriaziaren bidez tratatutako giltzurruneko litiasia, 1997an; ezkerreko zangoko zain sakoneko tronbosia, 2003ko urtarril eta maiatzean; hezueria bi oinetan; sorbaldetako tendinitisa eta pneumonia, 2012ko otsailan. **Ebakuntzak:** ezkerreko aldakaren eta **matrailaren** hausturak, trafiko-istripu batean, 1974an; disko-hernia eta odoluzkiak. **Ohiko tratamendua:** Valsartan, 160 mg-ko 1 pilula egunean; alopurinola, 300 mg-ko 1 pilula egunean; Sintrom, Hematologiako Zerbitzuak zehaztutako dosian.

**EGUNGO GAIXOTASUNA.** Aspaldiko traumatismoaren ondorioz, pazienteak **masailezurraren** eta tenporalaren arteko ankilosia zuen. 2013ko azaroaren 27an, programatutako ebakuntza egin zaio. Ankilosaturiko eremu osoa blokean erauzi eta alde zurretik prestaturiko plaka bat kokatu da hobi mandibularrean. Erauzketa egitean, kanpoko entzunbideko mintz tinpanikoa hautsi da. Otorrinologoko haustura baieztatu du eta neurri kontserbatzaileak eta jarraipen kliniko aholkatu ditu, miringoplastia baztertuz. Ebakuntza anestesia orokor balantzatuaren bidez egin da. Pazienteak ahoa zabaltzeko duen gaitasun txikia dela-eta, intubazio orotrakeala egin da fibroskopiaren bidez. Desintubazioa ebakuntza-gelan egin da konplikaziorik gabe.

**EBOLUZIOA BIZKORKETARAKO GELAN.** Kontziente dago, arnas autonomia du eta hemodinamikoki egonkor dago.

**ALTA EMATERAKOAN.** Tentsio arteriala, 100/50 mm Hg. Tratamendu basoaktiborik gabe. Bihotz-maiztasuna, 54 tau/min; O<sub>2</sub>Sat, % 98; arnas maiztasuna, 18 arnas/min. Kontziente eta orientatua dago perfusio onarekin. Aho-irekitze murriztua du. **BIHOTZAREN AUSKULTAZIOA.** Murmurio erritmikoak. **BIRIKEN AUSKULTAZIOA.** Arnas murmurioa entzuten da. Eskuin-birikaren behealde hipofonetiko dago eta ahoko bibrazioak ezabatuta daude **SABELALDEA.** Biguna eta zanpagarria; haztatzean mina sentitzen du hipogastrioan; ez dauka peritoneoko narritadurarik, sabel-soinu normalak. **GILTZURRUNAK.** Diuresi egokia, medikaziorik gabe.

**PROBA OSAGARRIAK. BOKIMIKA.** Kreatinina, 1,41 mg/dl. **HEMOGRAMA.** Hemoglobina, 12,3 g/dl; plaketak, 280.000; leukozitoak, 8.300. **KOAGULAZIOA.** Nazioarteko ratio normalizatua (INR), 1,1.

**DIAGNOSTIKOA TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ARTROPLASTIA. TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ANKILOSIA. BESTE AURREKARIAK.**

**TRATAMENDUA ETA GOMENDIOAK.** Anestesiako Zerbitzuaren aldetik ez du tratamendu edo jarraipen berezirik behar. Tratamendua, Ahoko eta Aurpegi-masailetakoa Kirurgiako medikuaren irizpidearen arabera.

Sinadura-  
data1/1  
orrialdea

Hona hemen Anestesiologia eta Berpiztea Osasun-txostenaren terminoen etiketatze emaitzak:

<p>✓ <b>OSPITALERATZEKO / KONTSULTARAKO ARRAZIOIA</b></p> <p>Masailezurraren eta tenporalaren arteko artroplastia egin ondoren, lekualdaturiko 67 urteko gizoneskoa.</p>
<p>✓ <b>ALTAREN ARRAZIOIA</b></p> <p>Aurpegi-masailetakoa eta Ahoko Kirurgiako Zerbitzura lekualdatzea.</p>
<p>✓ <b>AURREKARIAK</b></p> <p>Alergia die penizilinari, pobidona iododunari eta Anisakisari; erretzaille ohia da, hipertentsio arteriala du eta giltzurruneko gutxiegitasun arina; litotriaziaren bidez tratatutako giltzurruneko litiasia, 1997an; ezkerreko zangoko zain sakoneko tronbosia, 2003ko urtarril eta maiatzean; hezueria bi oinetan; sorbaldetako tendinitisa eta pneumonia, 2012ko otsailean. Ebakuntzak: ezkerreko aldakaren eta matrailaren hausturak, trafiko-istripa batean, 1974an; disko-hernia eta odoluzkiak. Ohiko tratamendua: Valsartan, 160 mg-ko 1 pilula egunean; alopurinola, 300 mg-ko 1 pilula egunean; Sintrom, Hematologiako Zerbitzuak zehaztutako dosian.</p>
<p>✓ <b>EGUNGO GAIXOTASUNA</b></p> <p>Aspaldiko traumatismoaren ondorioz, pazienteak masailezurraren eta tenporalaren arteko ankilosia zuen. 2013ko azaroaren 27an, programatutako ebakuntza egin zaió. Ankilosaturiko eremu osoa blokean erausi eta alde zurretik prestatutako plaka bat kokatu da hobi mandibularrean. Erauzketa egitean, kanpoko entzunbideko mintz tinpanikoa hautsi da. Otorrinolaringologoak haustura baieztatu du eta neurri kontserbatzaileak eta jarraipen klinikoa aholkatu ditu, miringoplastia baztertuz. Ebakuntza anestesia orokor balantzatuaren bidez egin da. Pazienteak ahoa zabaltzeko duen gaitasun txikia dela-eta, intubazio orotrakeala egin da fibroskopiaren bidez. Desintubazioa ebakuntza-gelan egin da komplikaziorik gabe.</p>

Masailezurraren eta tenporalaren arteko artroplastia egin ondoren, lekualdaturiko 67 urteko gizoneskoa.

Aurpegi-masailetakoa eta Ahoko Kirurgiako Zerbitzura lekualdatzea.

Alergia die penizilinari, pobidona iododunari eta Anisakisari; erretzaille ohia da hipertentsio arteriala du eta giltzurruneko gutxiegitasun arina; litotriaziaren bidez tratatutako giltzurruneko litiasia, 1997an; ezkerreko zangoko zain sakoneko tronbosia, 2003ko urtarril eta maiatzean; hezueria bi oinetan; sorbaldetako tendinitisa eta pneumonia, 2012ko otsailean. Ebakuntzak: ezkerreko aldakaren eta matrailaren hausturak, trafiko-istripa batean, 1974an; disko-hernia eta odoluzkiak. Ohiko tratamendua: Valsartan, 160 mg-ko 1 pilula egunean; alopurinola, 300 mg-ko 1 pilula egunean; Sintrom, Hematologiako Zerbitzuak zehaztutako dosian.

Aspaldiko traumatismoaren ondorioz, pazienteak masailezurraren eta tenporalaren arteko ankilosia zuen. 2013ko azaroaren 27an, programatutako ebakuntza egin zaió. Ankilosaturiko eremu osoa blokean erausi eta alde zurretik prestatutako plaka bat kokatu da hobi mandibularrean. Erauzketa egitean, kanpoko entzunbideko mintz tinpanikoa hautsi da. Otorrinolaringologoak haustura baieztatu du eta neurri kontserbatzaileak eta jarraipen klinikoa aholkatu ditu, miringoplastia baztertuz. Ebakuntza anestesia orokor balantzatuaren bidez egin da. Pazienteak ahoa zabaltzeko duen gaitasun txikia dela-eta, intubazio orotrakeala egin da fibroskopiaren bidez. Desintubazioa ebakuntza-gelan egin da komplikaziorik gabe.

Eranskina C. 5: Anestesiologia eta berpiztea osasun-txostenaren lehenengo zatia.

<p>✓ <b>EBOLUZIOA ETA IRUZKINAK BIZKORKETA-GELAN</b></p> <p>Kontziente dago, arnas autonomia du eta hemodinamikoki egonkor dago.</p>
<p>✓ <b>ALTA EMATERAKOAN</b></p> <p>Tentsio arteriala, 100/50 mm Hg. Tratamendu basoaktiborik gabe. Arnas maiztasuna, 18 arnas/min; bihotz-maiztasuna, 54 tau/min; O2Sat, % 98. Kontziente eta orientatua dago perfusio onarekin. Aho-irekitze murriztua du.</p> <p>✓ <b>Bihotzaren auskultazioa.</b> Murmurio erritmikoak</p> <p>✓ <b>Biriken auskultazioa.</b> Arnas murmurioa entzuten da. Eskuin-birikaren behealdea hipofonetiko dago eta ahoko bibrazioak ezabatuta daude</p> <p>✓ <b>Sabelaldea.</b> Bigua eta zanpagarria; hartzatzean mina sentitzen du hipogastrikoan; ez dauka peritoneoko narritadurarik, sabel-soinu normalak.</p> <p>✓ <b>Giltzurrunak.</b> Diuresi egokia, medikaziorik gabe.</p>

Kontziente dago, arnas autonomia du eta hemodinamikoki egonkor dago.

Tentsio arteriala, 100/50 mm Hg. Tratamendu basoaktiborik gabe. Arnas maiztasuna, 18 arnas/min; bihotz-maiztasuna, 54 tau/min; O2Sat, % 98. Kontziente eta orientatua dago perfusio onarekin. Aho-irekitze murriztua du.

Murmurio erritmikoak

Arnas murmurioa entzuten da. Eskuin-birikaren behealdea hipofonetiko dago eta ahoko bibrazioak ezabatuta daude

Bigua eta zanpagarria; hartzatzean mina sentitzen du hipogastrikoan; ez dauka peritoneoko narritadurarik, sabel-soinu normalak.

Diuresi egokia, medikaziorik gabe.

Eranskina C. 6: Anestesiologia eta berpiztea osasun-txostenaren lehenengo zatia.

<p>✓ <b>PROBA OSAGARRIEN LABURPENA</b></p> <p>✓ <b>Biokimika.</b> Kreatinina, 1,41 mg/dl.</p> <p>✓ <b>Hemograma.</b> Hemoglobina, 12,3 g/dl; plaketak, 280.000; leukozitoak, 8.300.</p> <p>✓ <b>Koagulazioa.</b> Nazioarteko ratio normalizatua (INR), 1,1.</p> <p>✓ <b>DIAGNOSTIKO NAGUSIA(K)</b> TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ARTROPLASTIA. TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ANKILOSIA. BESTE AURREKARIAK.</p> <p>✓ <b>TRATAMENDUA ETA GOMENDIOAK</b> Anestesiako Zerbitzuaren aldetik ez du tratamendu edo jarraipen berezirik behar. Tratamendua, Ahoko eta Aurpegi-masailetakoa Kirurgiako Medikuen irizpidearen arabera.</p>
--

Kreatinina, 1,41 mg/dl.

Hemoglobina, 12,3 g/dl; plaketak, 280.000; leukozitoak, 8.300.

Nazioarteko ratio normalizatua (INR), 1,1.

TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ARTROPLASTIA. TENPORALAREN ETA MATRAILEZURRAREN ARTEKO ANKILOSIA. BESTE AURREKARIAK.

Anestesiako Zerbitzuaren aldetik ez du tratamendu edo jarraipen berezirik behar. Tratamendua, Ahoko eta Aurpegi-masailetakoa Kirurgiako Medikuen irizpidearen arabera.

Eranskina C. 9: Anestesiologia eta berpiztea osasun-txostenaren azkeneko zatia.

# *D Eranskina*

---

---

## *Sprint-ak*

Gehienetan hiru-lau asteetan behin sprintak egingo dira. Hau da, eranskin honetan hiru (edo lau) asteko helburuak finkatuko dira (normalean ikasleak finkatuko ditu eta ahal duenean zuzendariekin batera) eta zeintzuk bete diren azalduko dira. Pakete bat bete ez bada, gorriz markatuta agertuko da eta honen zergatiak aipatuko dira. Gorriz markatatu izan den amaierako data, helburu horren benetako amaiera data agertuko da. Pakete bakoitzarekin dagokion atazaren zenbakia agertuko da.

GAP-aren entregatzeko epea hurbiltzen ari den heinean, *sprint*-ak motzagoak izango dira bilerak maizago egingo direlako, helburu berriak finkatuko direlako bileretan.

Hona hemen ataza bakoitzari dagokion zenbakia:

ATAZA IZENA	ZENBAKIA
Planifikazioa	1
Kudeaketa	2
Prestakuntza	3
Diseinua	4
Garapena	5
Probak	6
Memoria	7

Taula D. 1: *sprint*-ak gauzatzeko sortutako atazak.

Ataza hauek lan pakete txikietan banatuko dira azkar burutzeko. Ez dira behin betikoak, hau da, gerora ataza hau berriro beste *sprint* batean ager daiteke informazio gehigarria idatzi behar bada.

### *Lehenengo sprinta*

Denbora: 3 aste.

Hasiera Data: urtarrilak 25.

Bukaera Data: otsailak 15.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
SNOMED CT Started Guide Irakurri (3)	Otsailak 5	Otsailak 16	EZ	Testu zaila, hizkuntza nahiko tekniko eskaintzen duena (ingelesez), hainbat terminoen barneraketa
Klase diagrama (4)	Otsailak 14	Otsailak 15	BAI	-
VPN instalatu (3)	Otsailak 12	Otsailak 12	BAI	-
Python (3)	Otsailak 15	Otsailak 15	BAI	-

Taula D. 2: *lehenengo sprint*-a



### Bigarren sprinta

Denbora: 3 aste.

Hasiera Data: otsailak 15.

Bukaera Data: martxoak 7.

Finkatutako helburuak:

PAKETE(A) (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
SNOMED CT browser (3)	Otsailak 16	Otsailak 17	BAI	-
Kodea sakondu (3)	Martxoak 3	Martxoak 7	BAI	-
Aplikazioa martxan jartzen (3)	Martxoak 5	Martxoak 8	EZ	IXA taldeko zerbitzarietan baimenak falta ziren, aplikazioa martxan jarri ahal izateko
Memoriaren lehenengo atala (7)	Otsailak 18	Otsailak 18	BAI	-
SNOMED CT aurrekarietan idatzi (7)	Otsailak 19	Otsailak 24	BAI	-

Taula D. 3: bigarren sprint-a

### Hirugarren sprinta

Denbora: 3 aste.

Hasiera Data: martxoak 8.

Bukaera Data: martxoak 28.

Finkatutako helburuak:

PAKETE(A) (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
------------------------	--------------	--------------	----------	----------------------------

Zerbitzari – Bezero atala (7)	Martxoak 15	Martxoak 16	BAI	-
Ingeleseko bezeroa moldatu (5)	Martxoak 22	Martxoak 28	BAI	-
Teknologia atala sakondu (7)	Martxoak 16	Martxoak 18	BAI	-
Eskuliburua azpiatala (7)	Martxoak 16	Martxoak 16	BAI	-

Taula D. 4: hirugarren sprint-a

### Laugarren sprinta

Denbora: 1 aste.

Hasiera Data: martxoak 29.

Bukaera Data: apirilak 4.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Kalitatearen kudeaketa plana (1)	Martxoak 29	Martxoak 30	BAI	-
Komunikazio kudeaketa plana (1)	Martxoak 29	Martxoak 29	BAI	-
Arrisku kudeaketa plana (1)	Apirilak 1	Apirilak 1	BAI	-
LDE eta atazak finkatu (1)	Martxoak 29	Martxoak 29	BAI	-
Denboraren planifikazioa (1)	Martxoak 30	Martxoak 30	BAI	-

Mugarriaketa emangarriak (1)	Martxoak 31	Martxoak 31	BAI	-
------------------------------	-------------	-------------	-----	---

Taula D. 5: laugarren sprint-a

\*sprinta espero den baino lehenago amaituenez, sprint honen benetako amaierako data: **Apirilak 1**

### Bosgarren sprinta

Denbora: 4 aste.

Hasiera Data: apirilak 2.

Bukaera Data: apirilak 29.

Finkatutako helburuak:

PAKETEAK (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Lau objektu sortu klaseak.py-n (5)	Apirilak 2	Apirilak 5	BAI	-
Ingeleseko zerbitzaria moldatzen amaitu (5)	Apirilak 28	Apirilak 30	EZ	JSON-arekin hainbat arazo egon dira eta atal hau ezin izan da egin
Datu Basea (JSON) (5)	Apirilak 6	Apirilak 30	EZ	JSONa sortzeko arazoak eman ditu (deskribapen kopuru oso handia), beraz, denbora asko inbertitu da JSONa era egoki batean sortzeko (aste oso bat bi zerbitzariak martxan egon dira).
Memoria sakondu (7)	Apirilak 10	Apirilak 22	BAI	-

Taula D. 6: bosgarren sprint-a

### Seigarren sprinta

Denbora: 1 aste.

Hasiera Data: Apirilak 29 .

Bukaera Data: maiatzak 5.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Eustagger-aren informazio bilketa (7)	Maiatzak 2	Maiatzak 4	BAI	-

Taula D. 7: seigarren sprint-a

### Zazpigarren sprinta

Denbora: 1 aste.

Hasiera Data: maiatzak 5.

Bukaera Data: maiatzak 12.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Euskarazko zerbitzariaren lehenengo bertsioa (wraperra sortu) (5)	Maiatzak 7	Maiatzak 9	BAI	-
Memoria sakondu (7)	Maiatzak 9	Maiatzak 11	BAI	-

Taula D. 8: zazpigarren sprint-a

### Zortzigarren sprinta

Denbora: 1 aste.

Hasiera Data: maiatzak 12.

Bukaera Data: maiatzak 19.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
analizatailea_au.py sortu (5)	Maiatzak 12	Maiatzak 18	BAI	-
Memoria atal batzuk zuzendu (7)	Maiatzak 13	Maiatzak 18	BAI	-

Taula D. 9: zortzigarren sprint-a

### *Bederatzigarren sprinta*

Denbora: 1 aste.

Hasiera Data: maiatzak 19.

Bukaera Data: maiatzak 26.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
analisiak_au.js on sortu (5)	Maiatzak 19	Maiatzak 22	BAI	-
Analisi atala sakondu (7)	Maiatzak 21	Maiatzak 23	BAI	-
Inplementazio atala sakondu (7)	Maiatzak 22	Maiatzak 25	BAI	-
Osasun txosten ereduak (6)	Maiatzak 21	Ekainak 2	EZ	Eustagger_lite erremintak lema batzuek ez ditu ondo bueltatzen eta beraz ez du ondo

				deskribapenen ID ondo bilatzen.
--	--	--	--	---------------------------------

Taula D. 10: bederatzigarren sprint-a.

### Hamargarren sprinta

Denbora: 1 aste.

Hasiara Data: maiatzak 26.

Bukaera Data: ekainak 1.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Interfazea egin (5)	Maiatzak 26	Ekainak 8	EZ	Interfazea bukatzeko siuc zerbitzariari deitu behar zaio eta baimenen faltengatik ezin izan da honekin konexioa egin. AnaMed_eu-ekin konektatzea hainbat zailtasun sortu egin dira erabili den teknologiarengatik (JS, AJAX)
Bezeroa berrikusi hitz anitzerko terminoak bueltatzeko (5)	Maiatzak 26	Maiatzak 29	BAI	-
Inplementazio azpiatala sakondu (7)	Maiatzak 26	Ekainak 1	BAI	-

Taula D. 11: hamargarren sprint-a.

### *Hamabigarren sprinta*

Denbora: 1 aste.

Hasiera Data: ekainak 2.

Bukaera Data: ekainak 9.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Implementazio atala sakondu (7)	Ekainak 2	Ekainak 9	BAI	-
Jarraipen eta kudeaketa atala sakondu (2)	Ekainak 5	Ekainak 7	BAI	-

Taula D. 12: hamabigarren sprint-a.

### *Hamahirugarren sprinta*

Denbora: 1 aste.

Hasiera Data: ekainak 10.

Bukaera Data: ekainak 15.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Implementazio atala sakondu (7)	Ekainak 10	Ekainak 14	BAI	-
Analisiak_eu.js on fitxategi eguneratua sortu (5)	Ekainak 10	Ekainak 11	BAI	-
Teknologia azpiatala sakondu (7)	Ekainak 11	Ekainak 12	BAI	-

J+K sakondu (2)	Ekainak 11	Ekainak 13	BAI	-
Probak amaitu (6)	Ekainak12	Ekainak 12	BAI	-
Ondorioak atala (7)	Ekainak 13	Ekainak 14	BAI	-

Taula D. 13: hamahirugarren sprint-a

### *Hamalagarren sprinta*

Denbora: 4 egun.

Hasiera Data: ekainak 16.

Bukaera Data: ekainak 20.

Finkatutako helburuak:

PAKETE (ATAZA ZENB)	HASIERA DATA	BUKAERA DATA	AMAITUTA	DESBIDERAKETAREN ARRAZOIAK
Memoriaren zuzenketa (7)	Ekainak 17	Ekainak 20	BAI	-

Taula D. 14: hamalagarren sprint-a.



# *E Eranskina*

---

## Eskuliburuak

Bi eskuliburu sortu dira, `Med_en` sistema eta `OsateEUS` aplikazioa martxan jarri ahal izateko. Hona hemen:

### Med\_en-tzako eskuliburua

`Med_en` sistema martxan jartzeko terminalen beharra izango dugu. Hiru terminal beharko ditugu honetarako (bi zerbitzariantzako eta bestea bezeroarentzako).

Lehenik eta behin IXA taldeko kontua izan beharko dugu haien zerbitzarietan sartu ahal izateko. Hau lortu ondoren bi zerbitzariak (`corenlp.py` eta `ts_en.py`) martxan jarri beharko ditugu. Java-ren bertsioa aldatu beharko dugu aplikazioa exekutatu aurretik eta horretarako hurrengo bi agindu hauek exekutatu beharko ditugu:

```
setenv JAVA_HOME /usr/java/java18/
setenv PATH $JAVA_HOME/bin:$PATH
```

*Testu-kutxa E. 1: Java bertsioaren aldaketa*

Hau egin ondoren zerbitzariak martxan jarriko ditugu eta horretarako zerbitzariak dauden karpetara joango gara (KBP/zerbitzariak) eta hurrengo hau ikusiko dugu:

```
00-IRAKURRI corenlpython ts_en.py ts_eu.py
00-IRAKURRI~ __pycache__ ts_en.py~ util
```

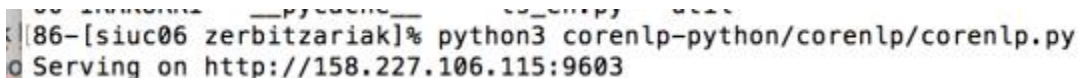
*Testu-kutxa E. 2: KBP/zerbitzari direktorioko edukia.*

Hemen gaudela lehenengo zerbitzaria martxan jarriko dugu (*Stanford*-ekoa, alegia) hurrengo hau idatziz:

```
python3 corenlpython/corenlpython/corenlpython.py
```

*Testu-kutxa E. 3: Stanford zerbitzaria martxan jartzen*

Dena ondo joan bada, hau ikusiko dugu terminalean:



```
186-[siuc06 zerbitzariak]% python3 corenlpython/corenlpython/corenlpython.py
Serving on http://158.227.106.115:9603
```

*Irudia E. 1: Stanford zerbitzaria martxan.*

Behin hau eginda, lehenengo zerbitzaria martxan izango dugu. Bigarrena (Terminologia Zerbitzaria, *TermMed\_eu*) martxan jartzeko karpeta berdinerara joango gara (`cd KBP/zerbitzariak`) eta `ts_en.py` kargatuko dugu (gogoratu beste terminal batean egin beharko dugula hau). Beraz:

```
python3 ts_en.py
```

*Testu-kutxa E. 4: : medTerm\_en zerbitzaria martxan jartzeko komandoa*

Zerbitzari hau martxan jartzeko, bi edo hiru minutu tartean itxaron beharko dugu dena kargatzeko (kontzeptuak, deskribapenak, erlazioak, ...). Dena ongi joan bada hurrengo hau ikusiko dugu:

```
/sc01a7/users/ixamed/BaliabideSemantikoak/SnomedCT_RF2Relea  
se_INT_20150731/Snapshot/Terminology/  
Language kargatuta 1087408  
Kontzeptuak kargatuta 317057  
Deskribapenak kargatuta 512724  
Erlazioak kargatuta 92807  
disorder finding-etik kenduta  
Hierarkiak kargatuta 317057  
Serving on http://158.227.106.115:9604
```

*Testu-kutxa E. 5: MedTerm\_en martxan*

Bi zerbitzariak martxan jarri ondoren analizatzailea jarri beharko dugu. Horretarako, analizatzaileak karpetara jo beharko dugu eta hau idatzi:

```
python3 analizatzailea_en.py "terminoa"
```

*Testu-kutxa E. 6: Terminoaren analisia egiteko komandoa.*

Non terminoa izango den erabiltzaileak sartuko duen terminoa (gaixotasuna). Terminoek beti kaketan bikoitzen artean (""") egongo da modu egoki batean lan egiteko. Adibidez:

```
73-[siuc06 analizatzaileak]% python3 analizatzailea_en.py  
"head ache"  
head ache  
[['head',  
  {'CharacterOffsetBegin': '0',  
   'CharacterOffsetEnd': '4',  
   'Hierarchy': 'BODYSTRUCTURE-0',  
   'Lemma': 'head',  
   'NamedEntityTag': 'O',  
   'PartOfSpeech': 'NN',  
   'sctId': '69536005#302548004'}]],  
 ['ache',  
  {'CharacterOffsetBegin': '5',  
   'CharacterOffsetEnd': '9',  
   'Lemma': 'ache',  
   'NamedEntityTag': 'O',  
   'PartOfSpeech': 'NN'}]]  
[['head', 'ache']]
```

*Testu-kutxa E. 7: "Head ache" terminoaren analisia.*

Hurrengo hau lortuko genuke eta bi zerbitzarietan hurrengo erantzunak:

ts\_en.py:

```
158.227.106.115 - - [16/Mar/2016 17:00:51] "POST / HTTP/1.1" 200 -  
158.227.106.115 - - [16/Mar/2016 17:00:51] "POST / HTTP/1.1" 200 -
```

*Testu-kutxa E. 8: MedTerm\_en jasotako erantzuna*

corenlp.py

```
head ache  
158.227.106.115 - - [16/Mar/2016 17:01:05] "POST / HTTP/1.1" 200 -
```

*Testu-kutxa E. 9: Stanford zerbitaria jasotako erantzuna*

## OsatEUS aplikazioaren eskuliburua

Aplikazioa atzitu ahal izateko erabiltzaileak hurrengo orrira jo beharko du: <http://siuc06.si.ehu.es:8000/login/>. Behin hau eginda *login* egiteko bi eremu agertuko zaigu:



Irudia E. 2: login egin leihoa

Hemen, administratzaileak sortu dituen erabiltzaileen “erabiltzaile izena” eta “pasahitza” era egokian sartu behar izango ditu. Aplikazio honetan administratzaileak soilik erabiltzaileak sor ditzake, gutxi batzuk erregistratuak egoteko. Erabiltzaileak ez badu bere pasahitza gogoratzen Pasahitza ahaztuta-ren gainean klik egingo du eta hurrengo leiho honetara joango da. Hemen egonda, bere *emaila* sartu behar izango du eta korreo bat iritsiko zaio bere pasahitza aldatzeko argibideekin.

### Berrezarri pasahitza

Pasahitza ahaztu duzu? Idatzi zure helbide elektronikoa eta berri bat ezartzeko jarraibideak bidaliko dizkizugu.

Helbide elektronikoa:

Berrezarri pasahitza

Irudia E. 3: Pasahitza berrezartzeko leihoa.

Aplikaziora sartu ondoren “Hasiera” deritzon leihoa ikusiko dugu hainbat espezialitate ezberdineko osasun txostenekin (Alta txosten berriak) eta jadanik idatzitan dauden altako osasun txostenekin.



Irudia E. 4: Erabiltzaile erregistratuaren hasierako leihoa.

Behin erabiltzeak erabakita zein alta txostenean idatzi nahi duen (Barne medikuntza, digestio-aparatua, pediatria, psikiatria, ...) hauetako batean gainean klik egin eta txostena osatzen joango da. Alta txosten bakoitza atal eta azpiatal jakin batzuk edukiko dituzte (hauen kudeaketa administratzaileak egingo ditu, 5.3 azpiatala). Hona hemen barne medikuntza espezialitatearen leihoa:



Irudia E. 5: Barne medikuntza alta txostena.

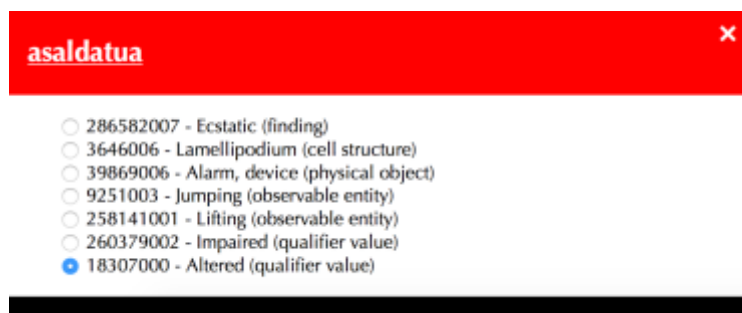
Ikus dezakegunez, erabiltzaileak bi atal bete beharko ditu eta ondoren “gorde” botoiaren gainean klik egin osasun-txosten hau gordetzeko. Hau egin aurretik, erabiltzailea atal batean edozer idazten badu eta ondoren kurtsorea idazteko eremutik kentzen badu, pantailaren eskuinaldean testu horren terminoak agertuko dira koloretan etiketatuak. Etiketa kolore bakoitza termino

mota bat adierazten du (4. Irudiko interfazearen goialdean agertzen dira termino mota erabilienak).

Ospitaleratzean medikazio antipsikotikoa jarri zaio . Solairuan bi aldiz asaldatu da .  
Tratamendua hasi eta hamar egunera mesfidantza desagertu zaio , baita eldarnioak eta  
haluzinazioak ere . Ez du gaixotasunaren kontzientziarik , eta botika hartzea  
gainbegiratu behar zaio , ezkututzen saiatzen baita , beraz ospitaleratzea luzatzea  
erabaki da egonaldi ertaineko ospitale batean . San Juan de\_Dios Ospitale Psikiatrikora  
lekualdatu da lau hilabeterako .

Irudia E. 6: terminoen detektatzea.

Koloreak dituzten hitzak (bakarra edo hitz anitzekoak) Med\_eu sistemak terminoak direla adierazten du. Termino hauek hitz anitzekoak edo bakunak izan daitezke. Gainera, azpimarratuta daudenak, termino anbiguoak direla adierazten dute, hau da, esanahi bat baino gehiago dutela. Behin hauen gainean klikatuta hauen *Full Specified Name* guztiak bueltatuko ditu ingelesez (euskarako FSN-ak ez daudelako prest oraindik).



Irudia E. 7: asaldatuta terminoaren PopUp leihoa.

Behin informazio hau guztia bistaratu ondoren “gorde” botoian klikatu eta sartutako informazioa ikusiko dugu beste orri batean. Hemen informazioa editatu ahal izango dugu (“Aldatu” *link*-aren bidez) edo atzera bota.

**Txostenaren laburpena**

Txosten\_id: 76

Editatu

**OSPITALERATZEKO ARRAZOIA:**  
Anemia duen 55 urteko gizonezkoa.

**AURREKARIAK:**  
Ez du alergiarik. Ez du ohiko tratamendurik.

[Aldatu](#) [Hasiera](#)

Irudia E. 8: barne medikuntza txostenaren laburpena.

Hau guztia azaldu ondoren, erabiltzaileak arazorik gabe aplikaziotik nabigatu ahal izango du.



# *F Eranskina*

---

---

## Kode zatiak

Eranskin honetan soilik kode zati handiak agertuko dira. Eranskin hau sortu da irakurlearen irakurketa errazago egiteko. Kode zati hauen erreferentziak 6. atalaren zehar egingo dira.

```

# terminoa .txt aldi baterako fitxategi batean gordeko da
def terminoTemp(self, term):
    filename = '/tmp/proba.txt'
    temp = open(filename, 'w+b')

    # terminoa utf-8 idazkeran berreskuratu
    terminoa = term.encode('utf-8')

    # aldi baterako fitxategi fitxategi idatzi
    temp.write(terminoa)
    temp.seek(0)

    return temp

#byte kate motatik String-era pasa eta erroa lortu
def zuhaitzBihurtu(self, naf):
    #byte motatik string motara pasa
    naf = naf[0].decode("utf-8")
    #print("naf", naf)

    #XMLaren erroa lortu
    root = ET.fromstring(naf)
    return root

```

*Testu-kutxa F. 2: terminoTemp eta zuhaitzBihurtu funtzioak*

```

def parse(self, term):
    jsonFinala = {}
    txtTemp = self.terminoTemp(term)
    #eustagger erabili datuak lortzeko
    eustagger = subprocess.Popen(['eustagger_lite', '-f',
'naf', '/tmp/proba.txt'], stdout=PIPE, close_fds=True)
    naf = eustagger.communicate()
    #erroa lortu string-etik
    root = self.zuhaitzBihurtu(naf)
    #JSON lortu
    json = self.naf2json(root)
    #gehitu 'sentences' sortzen ari garen JSONari
    jsonFinala.update({"sentences": json})
    #temp fitxategia itxi
    txtTemp.close()

    #analisiak_eu.json-ean idazteko
    self.pre_loaded_analysys_dict[term] = jsonFinala

    #10000 deskribapenero edo deskribapen guztiak lortu
    direnean fitxategian idatzi
    if len(self.pre_loaded_analysys_dict) % 10000 == 0 or
len(self.pre_loaded_analysys_dict) == 47051:
        self.storeAnalysis()

    return (jsonFinala)

```

*Testu-kutxa F. 1: parse funtzioa*

```

def naf2json(self, root):
    #wf guztiak lortu
    lista = []
    subList = []
    giltzak = {}
    #term eta wf etiketak lortu findall funtzioaren bitartez
    textEtiketa = root.findall("./text/wf")
    textTerm = root.findall("./terms/term")

    #etiketak hartu eta JSONean sartu...
    for i in range(len(textEtiketa)):
        case = []
        subList.insert(0, str(textEtiketa[i].text))
        #hurrengo atributuak lortu 'find'-en bidez eta giltzak listan gehitu
giltzak.update({'CharacterOffsetBegin':textEtiketa[i].find('.').attrib['offset']})
        giltzak.update({'CharacterOffsetEnd':
str(int(textEtiketa[i].find('.').attrib['length'])+
int(textEtiketa[i].find('.').attrib['offset'])-1)})
        giltzak.update({'Lemma':textTerm[i].attrib['lemma']})
        giltzak.update({'NamedEntityTag': 'O'})
        case = textTerm[i].attrib['case'].split()

        #case atribututik lehenengo bi laburtzapenak lortu
        if case[0] == 'IZE' or case[0] == 'ADJ' or case[0] == 'ADI' or case[0]
== 'LOT' or case[0] == 'ADB' or case[0] == 'DET' or case[0] == 'IOR':
            giltzak.update({'POS':case[0]})
            giltzak.update({'SubPOS':case[1]})
            case.pop(0)
            case.pop(0)
        elif case[0] == 'PRT' or case[0] == 'ITJ' or case[0] == 'BST':
            giltzak.update({'POS':case[0]})
            case.pop(0)

        caseEtiketak = self.laburtzapenak(case, giltzak, [])

        subList.insert(1, giltzak)

        lista.insert(i, subList)
        subList = []
        giltzak = {}
    return lista

```

*Testu-kutxa F. 3: naf2json funtzioa*

```

def laburtzapenak(self, case, caseGiltzak, funtzioSin):
    #case aldagaia hutsa dagoenean ([]) funtzioa amaitu
    if len(case) == 0:
        caseGiltzak.update({'funtzioSintaktikoak': funtzioSin})
        return caseGiltzak
    else:
        etiketa = []
        # @, < edo > hasten badira
        if '@' in case[0] or '<' in case[0] or '>' in case[0]:
            #funtzioSin-en sartu case[0]
            funtzioSin.insert(len(funtzioSin), case[0])
            case.pop(0)
            #errekurtsiboki deitu
            self.laburtzapenak(case, caseGiltzak, funtzioSin)
            return caseGiltzak
        else:
            #case[0] etiketari dagokion multzoa ezarri(BIZ, MUG, ORT, ...)
            etiketa = self.tokenEtiketak(case[0])
            #lehenego elementua ezabatu
            case.pop(0)
            caseGiltzak.update({etiketa[0]:etiketa[1]})
            #errekurtsiboki deitu
            self.laburtzapenak(case, caseGiltzak, funtzioSin)
            return caseGiltzak

```

*Testu-kutxa F. 4: laburtzapenak funtzioa*