

VILNIAUS GEDIMINO TECH UNIVERSITETAS

FUNDAMENTINIŲ MOKSLŲ FAKULTETAS

INFORMACINIŲ SISTEMŲ KATEDRA

Marius Šadreika

**SKRYDŽIŲ REKOMENDAVIMO SISTEMA ATSIŽVELGIANT Į ISTORINIŲ DUOMENŲ STATISTIKĄ**

**FLIGHTS RECOMMENDATION SYSTEM BASED ON STATISTICS OF HISTORICAL DATA**

Baigiamasis bakalauro darbas

Programų inžinerijos studijų programa, valstybinis kodas 612I30003

Programų sistemų studijų kryptis

Vilnius, 2021

**TURINYS**

[1 Lietuvos oro uostų statistika ir skrydžių paieškos sistemų nauda 6](#_Toc61132014)

[2 Panašių informacinių sistemų apžvalga 8](#_Toc61132015)

[2.1 „Expedia“ 8](#_Toc61132016)

[2.2 „Travelgenio“ 9](#_Toc61132017)

[2.3 „Skyscanner“ 11](#_Toc61132018)

[2.4 „Bravofly“ 12](#_Toc61132019)

[2.5 „Edreams“ 13](#_Toc61132020)

[2.6 „Opodo“ 15](#_Toc61132021)

[2.7 Apžvelgtų sistemų palyginimas 16](#_Toc61132022)

[3 Technologijų apžvalga 18](#_Toc61132023)

[3.1 „Back-end“ 18](#_Toc61132024)

[3.2 Duomenų bazė 20](#_Toc61132025)

[3.3 Duomenų ištraukimo veikimas 21](#_Toc61132026)

[3.4 Pagrindinės problemos: 22](#_Toc61132027)

[3.5 „Regex“ 22](#_Toc61132028)

[3.6 „.NET“ karkasas 22](#_Toc61132029)

[3.7 „Visual Studio“ 22](#_Toc61132030)

**Paveikslėlių sąrašas**

[1 pav. „Expedia“ grafinė vartotojo sąsaja 8](#_Toc61132031)

[2 pav. „Travelgenio“ grafinė vartotojo sąsaja 9](#_Toc61132032)

[3 pav. „Skyscanner“ grafinė vartotojo sąsaja 11](#_Toc61132033)

[4 pav. „Bravofly“ grafinė vartotojo sąsaja 12](#_Toc61132034)

[5 pav. „Edreams“ grafinė vartotojo sąsaja 13](#_Toc61132035)

[6 pav. „Opodo“ grafinė vartotojo sąsaja 15](#_Toc61132036)

[7 pav. Programavimo kalbų populiarumą tarp naudotojų atvaizduojantys grafikai 17](#_Toc61132037)

**Lentelių sąrašas**

[1 lentelė. 2019 ir 2020 įvykusių skrydžių statistika [2][3][4]. 6](#_Toc61132038)

[2 lentelė „Expedia“ tinklalapio SSGG analizė 8](#_Toc61132039)

[3 lentelė „Travelgenio“ tinklalapio SSGG analizė 10](#_Toc61132040)

[5 lentelė „Bravofly“ tinklalapio SSGG analizė 12](#_Toc61132041)

[6 lentelė „Edreams“ tinklalapio SSGG analizė 14](#_Toc61132042)

[7 lentelė „Opodo“ tinklalapio SSGG analizė 15](#_Toc61132043)

[8 lentelė Pasirinktų avialinijų palyginimas 16](#_Toc61132044)

**Įvadas**

Keliavimas – esamos lokacijos keitimas, kurį apibūdina: pirminė ir galutinė lokacijos, išvykimo ir atvykimo laikai, transporto priemonės. Keliavimas užima nemažą dalį mūsų gyvenimo, todėl atsižvelgus į kelionės tikslą atitinkamai pasirenkama transporto priemonė, kurios pagrindinės paskirtys yra šios: sutrumpinti, palengvinti kelionę ir užtikrinti saugumą. Viena iš pagrindinių transporto priemonių yra lėktuvas. Ši transporto priemonė padeda greitai ir saugiai įveikti didelius atstumus. Dėl šių savybių kelionės lėktuvu yra labai paklausios. Rinkoje šiuo metu yra nemažas pasirinkimas avialinijų, o jų teikiamų skrydžių išties daug. Neretai skrydžius, kurie atitinka asmens paieškos kriterijus, siūlo kelios avialinijos, o tai parodo stiprią konkurenciją tarp jų. Dėl šios priežasties avialinijos stengiasi pateikti geresnius skrydžių pasiūlymus nei jų konkurentai, nes neretai klientams svarbu, kad kelionės kaina ir jos trukmė būtų kuo mažesnė. Tam, kad avialinijos galėtų parengti geresnius pasiūlymus nei siūlo konkurentai, jie privalo stebėti rinką. Duomenų surinkimo, susisteminimo ir ataskaitų pateikimo sistema tampa labai svarbia sėkmingo konkurencingumo dalimi, nes būtent tokia sistema padeda stebėti rinką, o kadangi avialinijų ir jų pasiūlymų yra labai daug, rankiniu būdu surinkti visų duomenų beveik neįmanoma. Taip pat tokio tipo sistema būtų naudinga ne tik avialinijoms, bet ir įmonėms, kurios sudarinėja įvairius kelionių paketus. Sudarant kelionių paketą svarbu ne tik pateikti klientui geriausią nakvynės vietą, bet ir užtikrinti saugią, patogią, greitą ir rinkos atžvilgiu nebrangią kelionės kainą, nes nuo jos irgi priklauso kelionės paketo kaina, kuri klientui turi būti kuo tinkamesnė.

**Darbo objektas –** skrydžių rekomendavimo informacinė sistema.

**Darbo tikslas –** sukurti skrydžių rekomendavimo sistemą, kuri atsižvelgiant į kitų surinktų sistemų duomenis ir duomenų analizę, vartotojui pateiktų geriausią skrydžio pasiūlymą.

**Darbo uždaviniai:**

1. Išanalizuoti ir palyginti esamas skrydžių rekomendavimo sistemas, nustatyti pagrindinius trūkumus.
2. Atsižvelgiant į sistemos analogus ir jų trūkumus suformuoti kuriamos sistemos funkcinius ir nefunkcinius reikalavimus.
3. Remiantis sistemos reikalavimais sukurti veikiančios sistemos prototipą, kuriame būtų realizuoti suformuluoti reikalavimai.
4. Atlikti sukurtos sistemos testavimą ir pateikti galutines darbo išvadas.

# Lietuvos oro uostų statistika ir skrydžių paieškos sistemų nauda

Kiekvieną dieną atliekama daugybe skrydžių. Vien Lietuvoje 2019 metais Lietuvos oro uostai aptarnavo 6,5 mln. keleivių ir 62,4 tūkst. skrydžių. Lietuvoje yra 3 pagrindiniai oro uostai [1] :

1. Vilniaus oro uostas, kuris yra didžiausias civilinis oro uostas Lietuvoje. Jis 2019 metais aptarnavo 5 mln. keleivių.
2. Kauno oro uostas, kuris 2019 metais aptarnavo 1,1 mln. keleivių.
3. Palangos oro uostas, kuris 2019 metais aptarnavo 338 tūkst. keleivių.

**1 lentelė.** 2019 ir 2020 įvykusių skrydžių statistika [2][3][4].

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Vilniaus oro uostas | | | | | | | | | | | | |
| Metai | Sausis | Vasaris | Kovas | Balandis | Gegužė | Birželis | Liepa | Rugpjūtis | Rugsėjis | Spalis | Lapkritis | Gruodis |
| 2020 | 3567 | 3447 | 2232 | 353 | 498 | 976 | 1866 | 2274 | 1965 | 1517 | 1075 | 1087 |
| 2019 | 3501 | 3188 | 3565 | 4007 | 4471 | 4233 | 4165 | 4196 | 4324 | 4498 | 3697 | 3595 |
| Kauno oro uostas | | | | | | | | | | | | |
| Metai | Sausis | Vasaris | Kovas | Balandis | Gegužė | Birželis | Liepa | Rugpjūtis | Rugsėjis | Spalis | Lapkritis | Gruodis |
| 2020 | 779 | 724 | 599 | 248 | 289 | 453 | 635 | 693 | 566 | 495 | 302 | 373 |
| 2019 | 776 | 717 | 714 | 824 | 891 | 841 | 934 | 902 | 833 | 944 | 753 | 759 |
| Palangos oro uostas | | | | | | | | | | | | |
| Metai | Sausis | Vasaris | Kovas | Balandis | Gegužė | Birželis | Liepa | Rugpjūtis | Rugsėjis | Spalis | Lapkritis | Gruodis |
| 2020 | 355 | 379 | 187 | 6 | 72 | 206 | 373 | 438 | 423 | 244 | 164 | 102 |
| 2019 | 298 | 272 | 322 | 440 | 410 | 501 | 503 | 632 | 578 | 477 | 366 | 368 |

Skrydžių yra išties daug, o duomenų dar daugiau. Tų duomenų suvaldymas yra labai svarbus, nes tai padeda aviakompanijoms strategiškai konkuruoti. Būtent čia duomenų iš įvairių avialinijų surinkimo ir analizavimo bei palyginimo sistema gali praversti. Ja naudojantis naudotojas greitai gautų jam svarbiausią informaciją susijusią su jo nurodyta avialinija ar konkrečiu skrydžiu. Rinkoje jau yra panašaus tipo sistemų, kurios padeda avialinijoms suvokti rinkos esamą situaciją. Šio tipo sistemos yra reikalingos, nes be jų avialinijų paskirtam darbuotojui reikėtų rankiniu būdu jungtis prie avialinijos internetinio puslapio ir suvedinėti paieškos kriterijus tam, kad gautų reikiamą informaciją.

**Paieška standartiniame puslapyje:**

* Pradinio puslapio įsijungimas;
* Duomenų suvedimas (galima klaida, nes žmogus gali netyčia suvesti neteisingus duomenis ir tada atsirastų papildomas etapas – duomenų koregavimas);
* Konkretaus skrydžio internetiniame puslapyje užkrovimas;
* Pateiktų duomenų analizavimas;

**Būsimos sistemos privalumai:**

* Naudojantis būsima sistema pakartotinių jungimųsi ir pakartotinių duomenų suvedimų etapai bus praleidžiami.
* Duomenų filtravimą ir analizavimą atliks sistema, dėl to naudotojui nereikės pačiam apdoroti didelio srauto duomenų.
* Sistema išsaugos surinktus duomenis duomenų bazėje.

Tokio tipo sistema yra svarbi, nes ji padės suprasti rinkos pasiūlymus. Surinkti duomenys padės avialinijoms suplanuoti kainas atitinkamai pagal konkurentų kainas. Taip pat naudotojas galės suteikti tokius pasiūlymus, kurių konkurentai nesiūlo arba siūlo per mažai. Tokie duomenis taip pat padės susidaryti vaizdą, kurie skrydžiai yra populiariausi.

Šiuo metu panašių sistemų pagrindinė problema yra ta, kad nėra priėjimo prie jau įvykusių skrydžių. Tai apsunkina rinkos stebėjimą, nes sunku suprasti, ar to pačio skrydžio kaina kaip nors kito. Dėl to būsima sistema saugos praėjusių skrydžių duomenis. Tai reikalinga tam, kad naudotojas galėtų stebėti to pačio skrydžio, tos pačios avialinijos kainų šuolius.

# Panašių informacinių sistemų apžvalga

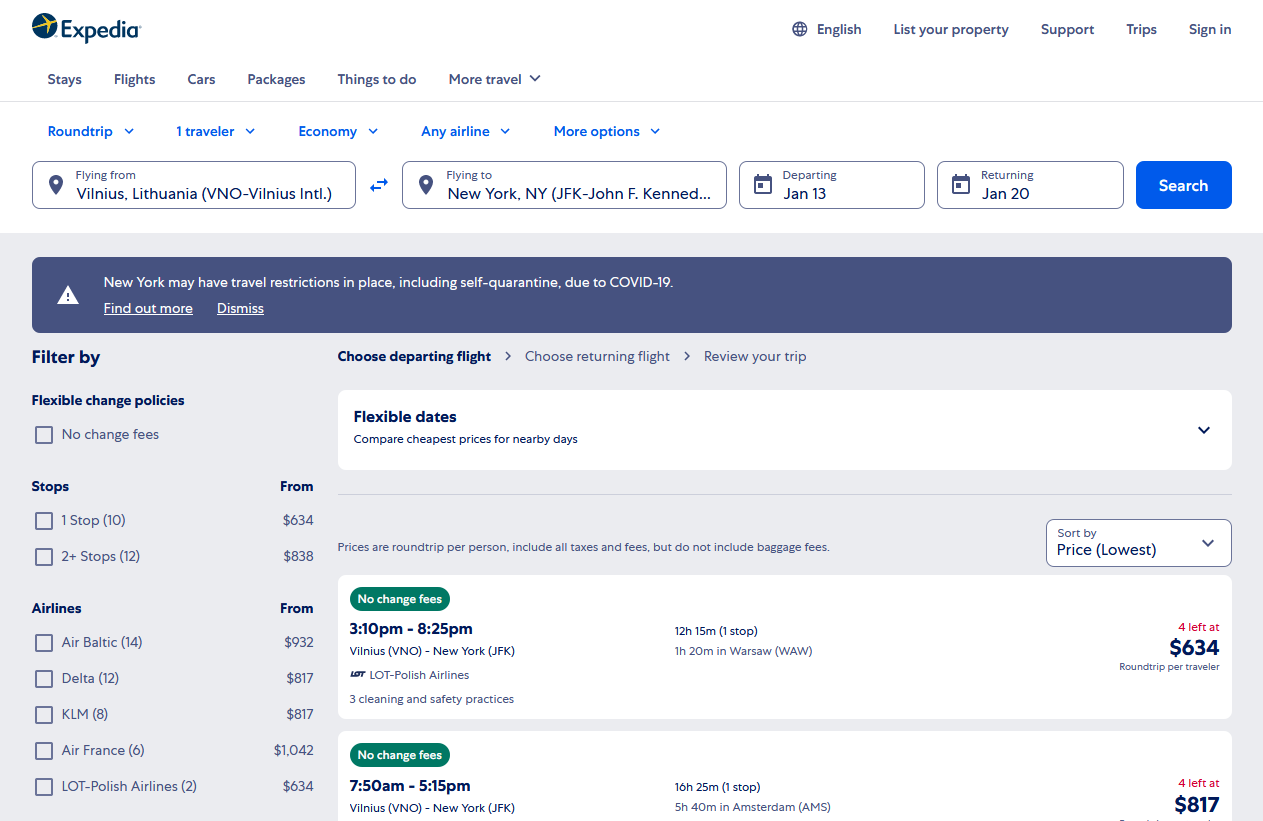
Stebėti siūlomų skrydžių rinką nėra paprasta. Paklausa ir pasiūla išties yra didelės, dėl to duomenų yra labai daug. Būtent dėl šios priežasties egzistuoja internetinės sistemos, kurios pagal naudojo užklausą pateikia skirtingų avialinijų pasiūlymus. Analitinės apžvalgos tikslas yra surinkti informacijos apie sistemas teikiančias informaciją apie skrydžius bei išanalizuoti problemas kylančias naudojantis šiomis sistemomis. Analizuojamiems tinklalapiams buvo keliamas tik šis reikalavimas: tinklalapyje turi būti pateikta informacija apie daugiau nei vieną avialiniją. Buvo analizuojamos populiariausios internetinės sistemos ir joms buvo atlikta SSGG (angl. *SWOT*) analizė.

SSGG (stiprybės, silpnybės, galimybės, grėsmės) arba *SWOT* (angl. *Strengths, Weaknesses, Opportunities, Threats*) analizė – metodas skirtas įvertinti organizacijos konkurencinę padėtį ir susidaryti strateginį planą. SSGG analizė vertina vidinius ir išorinius veiksnius, taip pat esamą ir būsimą potencialą [5][6].

* Stiprybės nurodo organizacijos pranašumą prieš konkurentus ir įmonės ypatybes, kurios teikia konkurencingumą rinkoje.
* Silpnybės nurodo tai, kas trukdo organizacijai dirbti optimaliai.
* Galimybės nurodo palankius išorinius veiksnius, kurie galėtų suteiki organizacijai konkurecinį pranašumą.
* Grėsmės nurodo veiksnius, galinčius pakenkti organizacijai.

## „Expedia“

„Expedia Inc.“ yra internetinė kelionių agentūra ir metaduomenų paieškos sistema, priklausanti Sietle įsikūrusiai amerikiečių bendrovei „Expedia Group“. Expedia.com buvo paleistas 1996 metais, spalio 22 d. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (1 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius skirtingų avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, pasirinkti sustojimų skaičių, filtruoti duomenis pagal kainą ar kelionės trukmę. Žemiau pateiktoje lentelėje (2 lentelė) yra pavaizduota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.



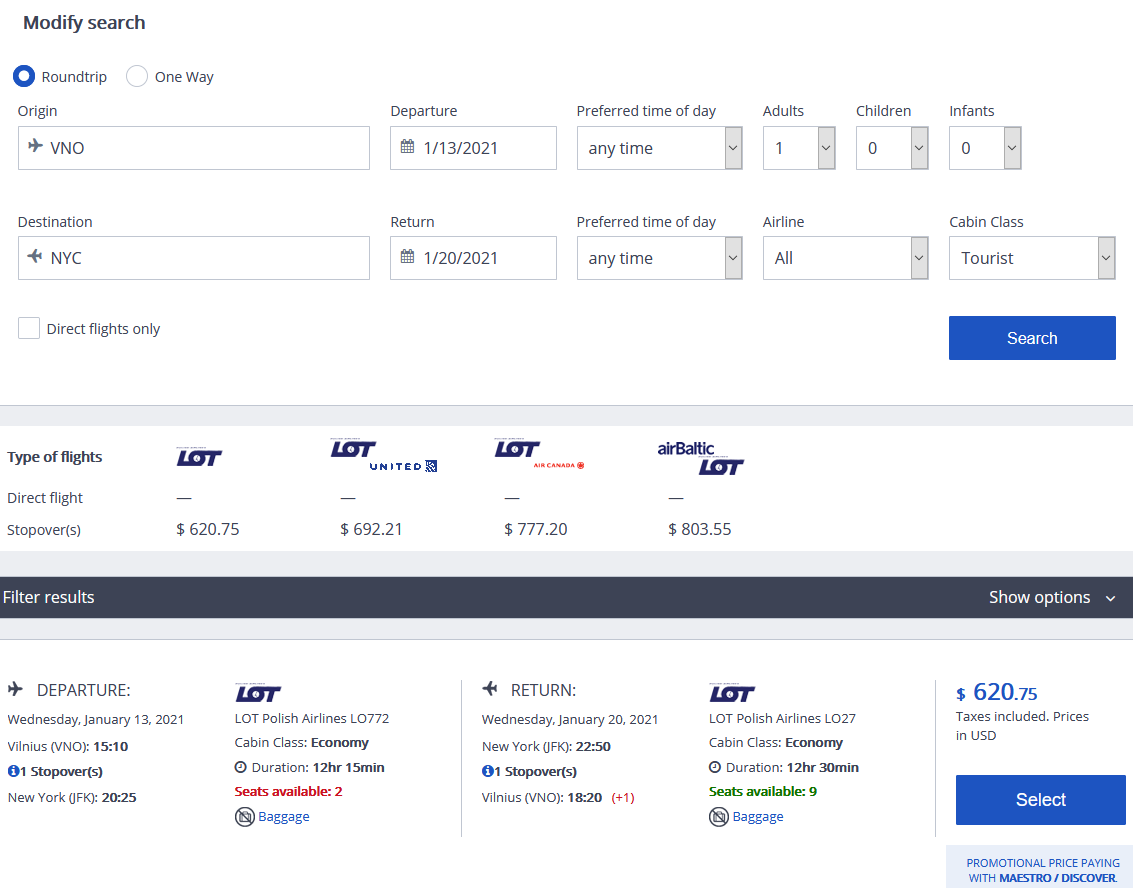
**1 pav.** „Expedia“ grafinė vartotojo sąsaja

**2 lentelė**. „Expedia“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Sustojimų skaičiaus pasirinkimo laukelius; 2. Skrydžių kiekio atvaizdavimą pagal sustojimo kartus; 3. Skrydžių kiekio atvaizdavimą pagal avialiniją; 4. Avialinijų mažiausių kainų atvaizdavimą; 5. Filtravimą pagal skrydžio trukmę ir kainą. | Šis internetinis puslapis neturi:   1. Tiesioginių skrydžių pasirinkimo paieškoje; 2. Kelionės trukmės pasirinkimo; 3. Skrydžio kainos intervalo pasirinkimo. |
| **Galimybės** | **Grėsmės** |
| Suteikti tiesioginių skrydžių pasirinkimo, kelionės trukmės, kainos intervalo pasirinkimo galimybes. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Expedia“ silpnybių. |

## „Travelgenio“

„Travelgenio“ yra internetinė kelionių agentūra, kuri buvo sukurta 2010 metais. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė, pageidaujamas dienos laikas, pageidaujama avialinija), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (2 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, filtruoti duomenis pagal kelionės trukmę. Žemiau pateiktoje lentelėje (3 lentelė) yra pavaizduota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.



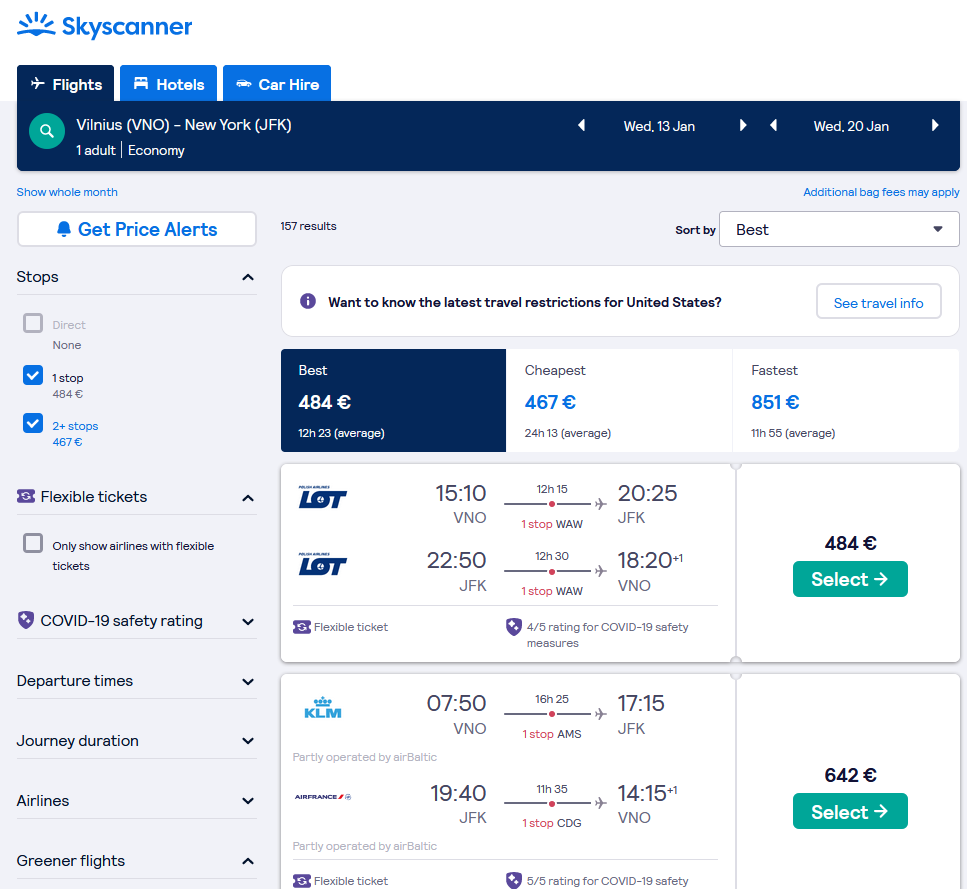
**2 pav**. „Travelgenio“ grafinė vartotojo sąsaja

**3 lentelė**. „Travelgenio“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Tiesioginių skrydžių pasirinkimą paieškoje; 2. Kelionės trukmės pasirinkimą; 3. Avialinijų mažiausių kainų atvaizdavimą. | Šis internetinis puslapis neturi:   1. Sustojimo skaičiaus pasirinkimo; 2. Skrydžių kiekio atvaizdavimo pagal sustojimo kartus; 3. Skrydžių kiekio atvaizdavimo pagal avialiniją; 4. Kainos intervalo pasirinkimo; 5. Filtravimo pagal skrydžio trukmę ir kainą. |
| **Galimybės** | **Grėsmės** |
| Suteikti sustojimo skaičiaus pasirinkimo, kainos intervalo pasirinkimo, filtravimo pagal skrydžio trukmę ir kainą galimybes. Atvaizduoti skrydžių kiekį pagal sustojimo kartus ir avialiniją. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Travelgenio“ silpnybių. |

## „Skyscanner“

„Skyscanner“ yra metaduomenų paieškos sistema ir kelionių agentūra, įsikūrusi Edinburge ir priklausanti didžiausiai Kinijos internetinių kelionių agentūrai „Trip.com Group“. Svetainė yra išversta daugiau nei į 30 kalbų. „Skyscanner“ buvo sukurtas 2002 metais. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė, suteikiamas tiesioginių skrydžių pasirinkimas), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (3 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, sustojimų skaičių, filtruoti duomenis pagal kelionės trukmę ir kainą. Žemiau pateiktoje lentelėje (4 lentelė) yra pavaidzuota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.



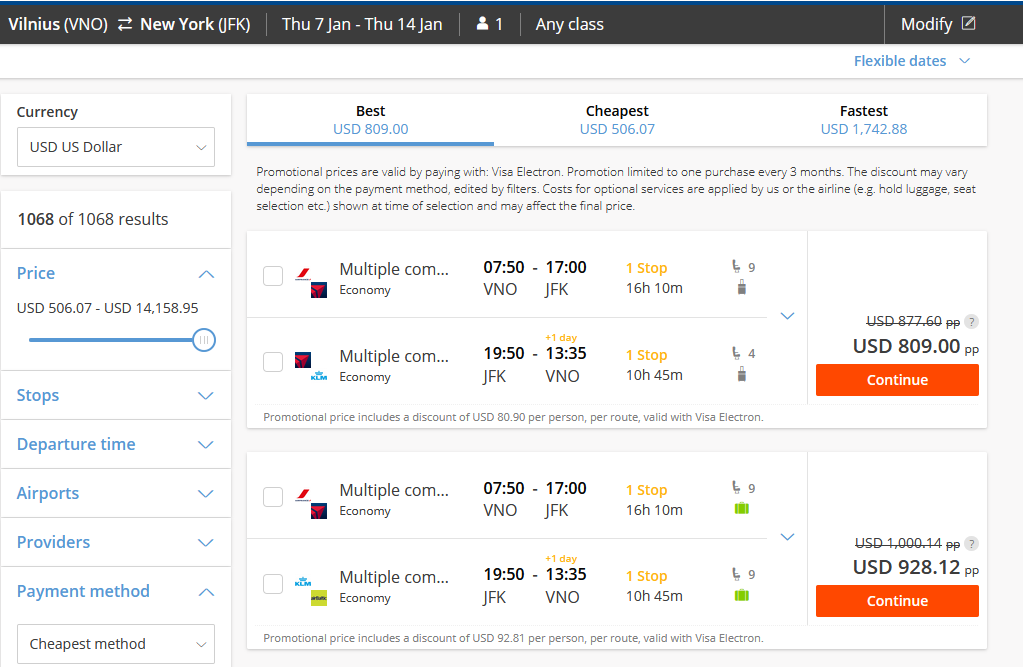
**3 pav.** „Skyscanner“ grafinė vartotojo sąsaja

**4 lentelė**. „Skyscanner“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Tiesioginių skrydžių pasirinkimą paieškoje; 2. Sustojimų skaičiaus pasirinkimą; 3. Kelionės trukmės pasirinkimą; 4. Filtravimą pagal skrydžio trukmę ir kainą. | Šis internetinis puslapis neturi:   1. Skrydžių kiekio atvaizdavimo pagal sustojimo kartus; 2. Skrydžių kiekio atvaizdavimo pagal avialiniją; 3. Kainos intervalo pasirinkimo; 4. Avialinijų mažiausių kainų atvaizdavimo. |
| **Galimybės** | **Grėsmės** |
| Suteikti kainos intervalo pasirinkimo galimybę. Atvaizduoti skrydžių kiekį pagal sustojimo kartus ir avialiniją. Atvaizduoti avialinijų mažiausias kainas. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Skyscanner“ silpnybių. |

## „Bravofly“

„Bravofly“ yra kelionių agentūra, kuri priklauso „Lastminute.com Group“. Svetainė yra išversta daugiau nei į 14 kalbų. „Bravofly“ buvo sukurtas 2006 metais. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (4 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, sustojimų skaičių, filtruoti duomenis pagal kelionės trukmę. Žemiau pateiktoje lentelėje (5 lentelė) yra pavaizduota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.



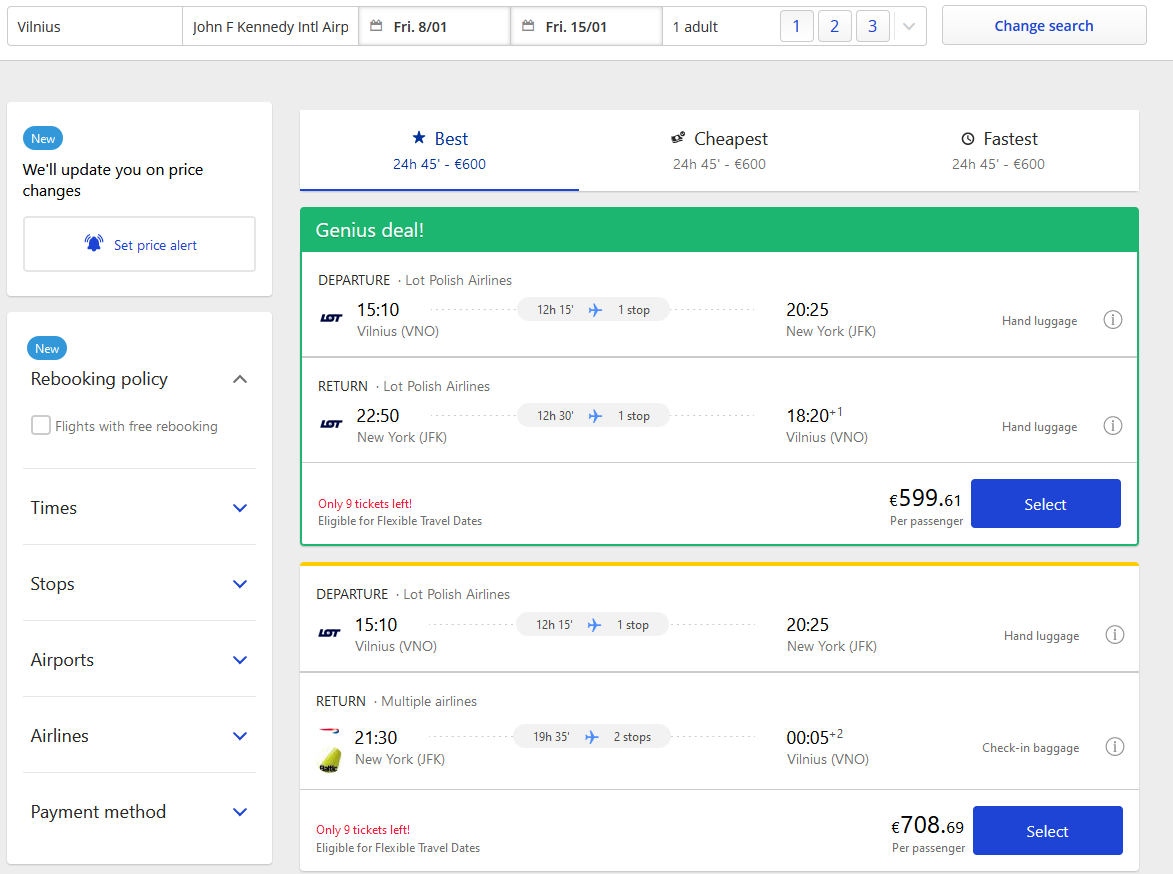
**4 pav**. „Bravofly“ grafinė vartotojo sąsaja

**5 lentelė.** „Bravofly“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Sustojimų skaičiaus pasirinkimą; 2. Kelionės trukmės pasirinkimą; 3. Kainos intervalo pasirinkimą. | Šis internetinis puslapis neturi:   1. Tiesioginių skrydžių pasirinkimo paieškoje; 2. Skrydžių kiekio atvaizdavimo pagal sustojimo kartus; 3. Skrydžių kiekio atvaizdavimo pagal avialiniją; 4. Avialinijų mažiausių kainų atvaizdavimo; 5. Filtravimo pagal skrydžio trukmę ir kainą. |
| **Galimybės** | **Grėsmės** |
| Suteikti tiesioginių skrydžių pasirinkimo, filtravimo pagal skrydžio trukmę ir kainą galimybes. Atvaizduoti skrydžių kiekį pagal sustojimo kartus ir avialiniją. Atvaizduoti avialinijų mažiausias kainas. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Bravofly“ silpnybių. |

## „Edreams“

„Edreams“ yra kelionių agentūra, kuri yra įsikūrusi Barselonoje, Ispanijoje. „Edreams“ buvo sukurtas 2000 metais. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (5 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, sustojimų skaičių, filtruoti duomenis pagal kelionės trukmę. Žemiau pateiktoje lentelėje (6 lentelė) yra pavaizduota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.

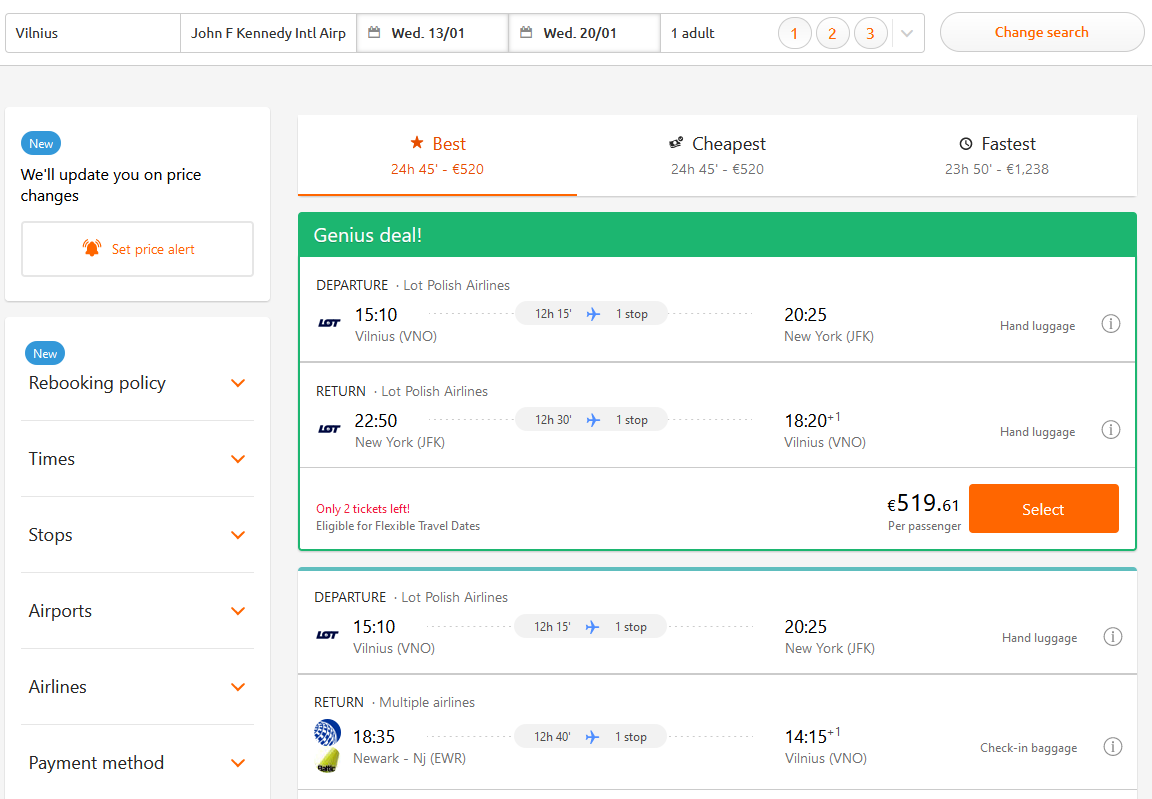


**5 pav.** „Edreams“ grafinė vartotojo sąsaja

**6 lentelė.** „Edreams“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Sustojimų skaičiaus pasirinkimą; 2. Kelionės trukmės pasirinkimą; 3. Filtravimą pagal skrydžio trukmę ir kainą. | Šis internetinis puslapis neturi:   1. Tiesioginių skrydžių pasirinkimo paieškoje; 2. Skrydžių kiekio atvaizdavimo pagal sustojimo kartus; 3. Skrydžių kiekio atvaizdavimo pagal avialiniją; 4. Avialinijų mažiausių kainų atvaizdavimo; 5. Kainos intervalo pasirinkimo. |
| **Galimybės** | **Grėsmės** |
| Suteikti kainos intervalo pasirinkimo, tiesioginių skrydžių pasirinkimo galimybes. Atvaizduoti skrydžių kiekį pagal sustojimo kartus ir avialiniją. Atvaizduoti avialinijų mažiausias kainas. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Edreams“ silpnybių. |

## „Opodo“

„Opodo“ yra kelionių agentūra, kurią įkūrė Europos oro linijos: „British Airways“, „Air France“, „Alitalia“, „Iberia“, „KLM“, „Lufthansa“, „Aer Lingus“, „Austrian Airlines“ ir „Finnair“ konsorciumas. „Opodo“ buvo įkurtas 2001 metais. Ši sistema naudotojui pateikia paieškos laukus (išvykimo ir atvykimo lokacija, išvykimo ir atvykimo datos, kelionės tipas, klasė), kuriuos užpildžius atidaromas pagrindinis informacijos atvaizdavimo langas (6 paveiksliukas). Sistema atvaizduoja užklausą atitinkančius avialinijų skrydžių pasiūlymus. Informacijos atvaizdavimo lange taip pat yra pateikiamos galimybės dar labiau patikslinti poreikius: galima nurodyti konkrečią avialiniją, sustojimų skaičių, filtruoti duomenis pagal kelionės trukmę. Žemiau pateiktoje lentelėje (7 lentelė) yra pavaizduota SSGG internetinės sistemos analizė, joje atvaizduojama sistemos stipriosios ir silpnosios pusės.

**6 pav**. „Opodo“ grafinė vartotojo sąsaja

**7 lentelė.** „Opodo“ tinklalapio SSGG analizė

|  |  |
| --- | --- |
| **Stiprybės** | **Silpnybės** |
| Šis internetinis puslapis turi:   1. Tiesioginių skrydžių pasirinkimą paieškoje; 2. Sustojimų skaičiaus pasirinkimą; 3. Kelionės trukmės pasirinkimą; 4. Filtravimą pagal skrydžio trukmę ir kainą. | Šis internetinis puslapis neturi:   1. Skrydžių kiekio atvaizdavimo pagal sustojimo kartus; 2. Skrydžių kiekio atvaizdavimo pagal avialiniją; 3. Kainos intervalo pasirinkimo; 4. Avialinijų mažiausių kainų atvaizdavimo. |
| **Galimybės** | **Grėsmės** |
| Suteikti kainos intervalo pasirinkimo, galimybę. Atvaizduoti skrydžių kiekį pagal sustojimo kartus ir avialiniją. Atvaizduoti avialinijų mažiausias kainas. | Konkurentų internetinis puslapis gali būti patogesnis dėl „Opodo“ silpnybių. |

## Apžvelgtų sistemų palyginimas

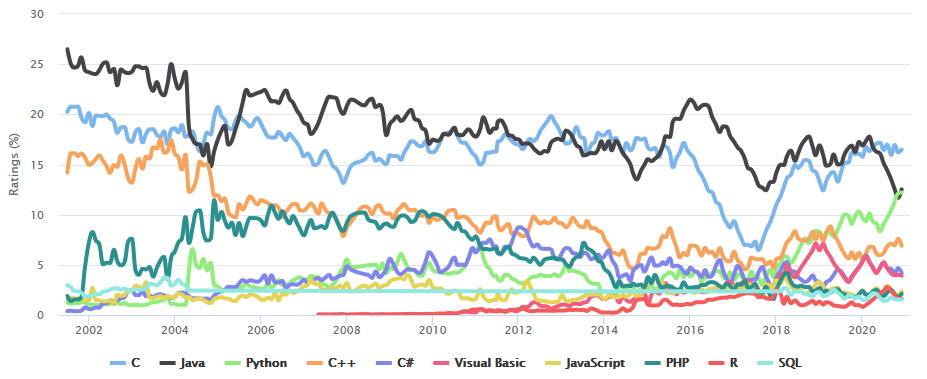
Išvykimo ir atvykimo lokacijų, atvykimo ir išvykimo laikų, kelionės tipo, skrydžio klasės ir avialinijos pasirinkimai yra visuose šiuose internetiniuose puslapiuose. Skirtumai tarp internetinių puslapių pateikti lentelėje.

**8 lentelė.** Pasirinktų avialinijų palyginimas

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Pavadinimas** | „Expedia“ | „Travelgenio“ | „Skyscanner“ | „Bravofly“ | „Edreams“ | „Opodo“ |
| **Tiesioginių skrydžių pasirinkimas paieškos laukelyje** | Nėra | Yra | Yra | Nėra | Nėra | Yra |
| **Sustojimų skaičiaus pasirinkimas** | Yra | Nėra | Yra | Yra | Yra | Yra |
| **Skrydžių skaičiaus atvaizdavimas pagal sustojimo kartus** | Yra | Nėra | Nėra | Nėra | Nėra | Nėra |
| **Kelionės trukmės pasirinikimas** | Nėra | Yra | Yra | Yra | Yra | Yra |
| **Skrydžių kiekio atvaizdavimas pagal avaialiniją** | Yra | Nėra | Nėra | Nėra | Nėra | Nėra |
| **Mažiausios ir didžiausios kainų pasirinkimas** | Nėra | Nėra | Nėra | Yra | Nėra | Nėra |
| **Avialinijų mažiausių kainų atvaizdavimas** | Yra | Yra | Nėra | Nėra | Nėra | Nėra |
| **Filtravimas pagal skrydžio trukmę, kainą** | Yra | Nėra | Yra | Nėra | Yra | Yra |

# Technologijų apžvalga

Kuriant skrydžių rekomendavimo informacinę sistemą svarbu, kad pasirinkti įrankiai būtų atidžiai išanalizuoti ir pasirinkti tik tinkamiausi. Vienas iš pagrindinių įrankių, kuriuos reikia apsvarstyti, yra programavimo kalba. Programavimo kalbų pasirinkimas yra išties didelis, bet svarbu paminėti ir tai, kad skirtingos programavimo kalbos turi skirtingus pranašumus ir trūkumus. Atsižvelgus į šiuos programavimo kalbų aspektus galima lengviau išsirinkti labiausiai tinkantį variantą. Taip pat šį sprendimą priimant gali padėti populiariausių programavimo kalbų statistika. Remiantis „TIOBE“ kokybės indikatoriumi, kur vertinimas atliekamas remiantis ISO 25010 PĮ kokybės standartu sudarytas grafikas, kuris atvaizduoja programavimo kalbų reitingus nuo 2002 metų iki 2020 metų. ISO 25010 standartas vertina PĮ iš įvairių perspektyvų: įsigijimo, reikalavimų, kūrimo, naudojimo, vertinimo, palaikymo, priežiūros, kokybės užtikrinimo ir audito [7].



7 pav. Programavimo kalbų populiarumą tarp naudotojų atvaizduojantys grafikai

Atsižvelgus į įsigijimo, reikalavimų, kūrimo, naudojimo, vertinimo, palaikymo, priežiūros, kokybės užtikrinimo ir audito kriterijus matoma, kad remiantis „TIOBE“ kokybės indikatoriumi verta apsvarstyti šias pirmąsias penkias programavimo kalbas: „C“, „Java“, „Python“, „C++“, „C#“.

## „Back-end“

„C“ struktūrinė programavimo kalba buvo sukurta 1972 metais. Pradžioje ši programavimo kalba buvo sukurta kaip sistemos programavimo kalba, kuri skirta operacinėms sistemos kurti. Labiausia ši programavimo kalba tinka aparatinės įrangos programoms, mikroschemų dizaino kūrimui. Pagrindiniai privalumai „C“ programavimo kalbos: žemo lygio prieiga prie atminties, turi paprastą raktinių žodžių rinkinį [8].

„C++“ į objektą orientuota programavimo kalba, kuri buvo sukurta 1985 metais kaip „C“ programavimo kalbos plėtinys [9]. „C++“ lyginant su „C“ turi daug geresnę duomenų apsaugą. Taip pat „C“ programavimo kalba palaiko tik primityvius, iš anksto apibrėžtus duomenų tipus. Į tuos duomenų tipus nėra įtraukti loginiai (angl. *boolean*) arba eilutės (angl. *string*) duomenų tipai. „C++“ programavimo kalba leidžia kurti net savo apibrėžtus duomenų tipus. „C++“ palaiko operatoriu perkrovas (angl. *operator overloading*) ir funkcijų perkrovas (angl. *function overloading*), o tai suteikia programuotojams galimybę sukurti savo operatorius ir funkcijas. „C++“ taip pat naudoja tiesiogines funkcijas (angl. *inline functions*) organizavimui, bandymo ir pagavimo blokus (angl. *try and catch blocks*) klaidų manipuliavimui. „C“ programavimo kalba neturi nei vieno iš šitų išvardintų galimybių. [10] Palyginus šias dvi programavimo kalbas šiai sistemai kurti būtų tinkamesnė „C++“ programavimo kalba .

„C#“ į objektą orientuota programavimo kalba, kuri buvo sukurta 2000 metais. „C#” programavimo kalba buvo sukurta įmonės „Microsoft Corporation“. „C#” yra aukštos abstrakcijos programavimo kalba (angl. *high level of abstraction*). Tai reiškia, kad ši programavimo kalba yra labiau priimtina žmogui dėl savo paprastos sintaksės. Šiuo atveju „C#” yra pranašesnė negu „C++“, nes „C++“ abstrakcijos lygis yra pakankamai žemas. Taip pat „C#” turi angl. *garbage collection*. [11] Angl. *garbage collection* – automatinis atminties valdymas, tai reiškia, kad jeigu atmintį užima objektai, kurie nėra naudojami, jie yra automatiškai pašalinami. „C++” šio funkcionalumo neturi, dėl to atminties valdymą reikia atlikti patiems, o tai reiškia, kad kodo užrašyma tampa sudėtingesnis, nes reikia galvoti ne tik apie problemą, kurią sukurta sistema turėsi išspręsti, bet ir apie taip, kaip teisingai valdyti atmintį. „C++” programavimo kalba leidžia atlikti daugelį veiksmų iki tol kol sintaksė išlieka teisinga dėl tokio programavimo kalbos lankstumo gali atsirasti ir tam tikrų netikėtų operacinės sistemos pažeidimų. „C#” yra daug labiau apsaugota ir pateikia kompiliatoriui klaidų ir įspėjimų pranešimus, kurie apsaugo nuo rimtų klaidų, kurias „C++” gali leisti. Taip pat „C#” turi „LINQ“ (angl. *Language Integrated Query*), kuris yra labai svarbus dirbant su dideliu duomenų srautu [12]. Atsižvelgus į išvardintus esminius kriterijus būsimai sistemai kurti būtų naudingesnė „C#” programavimo kalba negu „C“ ar „C++“ programavimo kalbos.

„Python” yra interpretuojama, aukšto abstrakcijos lygio, į objektą orientuota programavimo kalba, kuri buvo sukurta 1991 metais. Kadangi „Python“ yra interpretuojama programavimo kalba, prieš paleidžiant kodą nereikia jo kompiliuoti . „Python“ turi „Python Package Index“ (PyPI) - daugybė trečiųjų šalių modulių, leidžiančių „Python“ sąveikauti su daugeliu kitų kalbų ir platformų. „Python“ taip pat teikia didelę biblioteką, kuri turi net interneto protokolus, eilutės (angl. *String*) operacijas, žiniatinklio paslaugų įrankius ir operacinės sistemos sąsajas [13][14]. Bibliotekoje jau yra dažno naudojimo programavimo užduotys, kurios gali sumažinti kodo ilgį. Kadangi „C#” programavimo kalba gali atlikti beveik viską, ką geba atlikti „Python“ programavimo kalba, bet „C#” struktūra yra aiškesnė ir labiau organizuota, o veikimo laikas (angl. *runtime*) yra trumpesnis, dėl to „C#” vis dar būtų tinkamesnis pasirinkimas.

„Java“ yra į objektą orientuota programavimo kalba, kuri buvo sukurta 1995 metais. „Java“ gali būti realizuota bet kokioje platformoje. „Java“ veikimas yra labai greitas, bet sunaudojama daug atminties, nes ši programavimo kalba yra kompiliuojama. „Java“ programavimo kalbą pakankamai lengva išmokti. [15][16] Lyginant „Java“ ir „C#” abi programavimo kalbos yra tikrai geri pasirinkimai.

## Duomenų bazė

Yra dviejų tipų duomenų bazės: SQL ir NoSQL. SQL duomenų bazė yra lentelėmis paremta duomenų bazė, o NoSQL dokumentais, raktų poromis, grafais. SQL duomenų bazių veikimo principas yra struktūrizuotos duomenys lentelėse, kuriuose kiekviena eilutė yra duomuo, o stulpeliuose reikšmės ir pagrindinis skirtumas tarp NoSQL yra ryšiai tarp lentelių [17][18][19].

„Microsoft SQL Server“ – reliacinė duomenų bazių valdymo sistema, kuri palaiko platų transakcijų procesų, verslo žinių ir analitinių programų pasirinkimą įvairiose informacinių technologijų aplinkose [20][21]. „Microsoft SQL Management Studio“ – integruota „SQL Server“ aplinka infrastruktūros valdymui. „MSSQL Management Studio“ turi naudotojo sąsają ir įrankių grupę su plačiu skriptų redaktorių pasirinkimu, kurie sąveikauja su „SQL Server.

„MongoDB“ – yra šiuo metu populiariausia NoSQL duomenų bazė, nes duomenis saugo JSON formatu kaip dokumentus, pasižymi dideliu duomenų apdorojimo greičiu, nes nėra ryšių tarp lentelių, kas sulėtina duomenų apdorojimą labai sudėtingoje lentelėje . Lengvai prižiūrima ir administruojama [22].

Kadangi „Microsoft SQL Server“ yra palankesnis pasirinkimas dirbant su „C#” programavimo kalba naudojant „.NET” karkasą, bus pasirinkta šis įrankis. „.NET” karkasas bus naudojamas dėl savo skaitomumo, saugumo ypatybių. Taip pat grafinei naudotojo sąsajai pateikiamas „Windows Forms“ pasirinkimas, kurį lengva valdyti ir puikiai tinka būsimo projekto grafinės naudotojo sąsajos realizavimui. [23]

## Duomenų ištraukimo veikimas

Pasaulinis tinklas (angl. *World Wide Web*) – informacijos gavybos paslauga teikiama internete. Internetas – pasaulinis kompiuterių tinklas. Pasaulinio tinklo paslauga buvo sukurta Sir Tim Berners 1989 metais. Taip pat Sir Tim Berners 1990 metais sukūrė 3 fundamentines technologijas, kurios palaiko kompiuterių komunikavimą [24].

* *HTML* (angl. *HyperText Markup Language*) – tinklalapių kodavimo sistema, skirta suformatuoti ir atvaizduoti duomenis internete [25]. *HTML* aprašomas tekstiniame faile, kuriame naudojamos žymos (angl. *tags*), kurios ir padeda sustruktūrizuoti internetinį puslapį.
* *URI* (angl. *Uniform Resource Identifier*) – identifikatorius, kuris naudojamas jungiantis prie internetinio puslapio [26].
* *HTTP* (angl. *Hypertext Transfer Protocol*) – programinio lygmens protokolas skirtas: bendradarbiavimo, hipermedijos informacijos sistemoms [27].

Duomenų ištraukimas robotas (angl. *data scraper* arba angl. *web crawler*) – tai programa, programinė įranga arba programinis kodas (angl. *programmed script*), kuris atlieka užklausas (ieško pausauliniame tinkle) ir siunčiasi žiniatinklio dokumentus, iš kurių vėliau ištraukiami reikalingi konkretūs duomenys. 1994 metais buvo sukurtas pirmasis duomenų ištraukimo robotas [28].

Programos skirtos duomenų ištraukimui neretai vadinamos robotais. Robotų paskirtis ištraukti reikalingus duomenis. Gero duomenų ištraukimo roboto savybė – robotas neturi apkrauti internetinio puslapio nereikalingomis užklausomis. Kuo daugiau užklausų atlieka robotas, tuo didesnis yra šansas, kad puslapis pradės blokuoti kitas roboto užklausas, o tai reiškia, kad robotas nebegalės ištraukti duomenų. Tam, kad puslapis nesuvoktų, kad į jį kreipiasi robotas, robote yra aprašomi naršyklės parametrai, taip imituojama naršyklės užklausa, kai asmuo jungiasi prie puslapio.

## Pagrindinės problemos:

1. Blokavimas, nes puslapis neleidžia robotui prisijungti, o tai reiškia, kad duomenys būna neištraukti.
2. Internetinio puslapio atnaujinimai, nes duomenys ištraukiami iš internetinio puslapio *HTML* arba *JSON* formato failo, kuris kraunamas su užklausa. Jeigu duomenys yra ištraukiami iš *HTML*, naudojama „Regex“, kuris padeda šiek tiek užtrikrinti duomenų ištraukimą, kai *HTM*L pakinta. Jeigu duomenys yra traukiami iš JSON, šanas, kad duomenys bus neištraukti yra mažesnis. Roboto taisymas taip pat lengvesnis. Nors robotas gali atdaptuotis prie mažų pokyčių, dideli pokyčiai internetiniame puslapyje, iš kurio ištraukiami duomenys, gali privesti prie duomenų neištraukimo.

Duomenų išgavimo procesas susidaro kiekvienam internetiniam puslapiui yra vis kitoks. Internetiniai puslapiai turi skirtingas struktūras ir apsisaugojimus.

## „Regex“

„Regex“ (angl. *Regular expression*) – įrankis, kuris pagal nurodytą modelį atlieka paieška nurodytame tekste. Modelis (angl. *pattern*) – simbolių seka, kuri nurodo, kokių duomenų tekste reikia ieškoti. Regex turi savo sutartinių simbolių biblioteką, kuri palengvina darbą ir leidžia tekste surasti ne tik konkrečią informaciją, bet surasti didelius informacijos blokus, kurie tenkina nurodytas savybes [29].

## „.NET“ karkasas

„.NET“ karkasas yra programų kūrimo platforma sukurta „Microsoft“ 2002 metais, kuri yra skirta „Windows“ programų kūrimui ir paleidimui. „.NET“ karkasas turi programų kūrimo įrankius, programavimo kalbas ir bibliotekas, skirtas kurti darbalaukio, žiniatinklio programas, žaidimus. Šis karkasas palaiko daugiau negu 60 programavimo kalbų, iš kurių 11 programavimo kalbų yra sukurtos „Microsoft“ [30].

## „Visual Studio“

„Visual Studio“ yra integruota programų kūrimo aplinka, kuri būvo pristatyta 1997 metais įmonės „Microsoft“. „Visual Studio“ turi kodo redaktorių, kuris palaiko „IntelliSense“.

„Visual Studio“ palaiko 36 skirtingas programavimo kalbas [31]. Derintuvas (angl. *debugger*) veikia tiek kodo, tiek mašininame lygmenyje. Derintuvas taip pat teikia angl. *breakpoints* galimybę, kuri leidžia sustabdyti programą tam tikroje vietoje. Tokia galimybė yra naudinga norint įsitikinti ar kodas fragmentas veikia korektiškai.

**Literatūros šaltiniai**

[1] Lietuvos oro uostai „Apie Lietuvos oro uostus“ 2020 [žiūrėta 09 01 2021]. Prieiga per internetą: <https://www.ltou.lt/lt/apie-lietuvos-oro-uostus/apie-lietuvos-oro-uostus-1>

[2] Lietuvos oro uostai „VNO“ 2020 [žiūrėta 09 01 2021]. Prieiga per internetą: <https://www.ltou.lt/lt/aviacines-paslaugos/skrydziu-statistika/vno-2>

[3] Lietuvos oro uostai „KUN“ 2020 [žiūrėta 09 01 2021]. Prieiga per internetą: <https://www.ltou.lt/lt/aviacines-paslaugos/skrydziu-statistika/kun-2>

[4] Lietuvos oro uostai „PLQ“ 2020 [žiūrėta 09 01 2021]. Prieiga per internetą: <https://www.ltou.lt/lt/aviacines-paslaugos/skrydziu-statistika/plq-2>

[5] „MindTools“ „SWOT Analysis“ [žiūrėta 09 01 2021] Prieiga per internetą: <https://www.mindtools.com/pages/article/newTMC_05.htm>

[6] M. Grant „Strenght, Weakness, Opportunity, and Threat(SWOT) Analysis 2020[žiūrėta 09 01 2021] Prieiga per internetą: <https://www.investopedia.com/terms/s/swot.asp>

[7] „Tiobe“ the software quality company 2021[žiūrėta 07 12 2020] Prieiga per internetą: <https://www.tiobe.com/tiobe-index/>

[8] The Editors of Encyclopeadia Britannica „C“ 2017 [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.britannica.com/technology/C-computer-programming-language>

[9] V. Singh „Difference Between C and C+“ 2020 [žiūrėta 07 12 2020] Prieiga per internetą: <https://hackr.io/blog/difference-between-c-and-cplusplus>

[10] Laura M. „C VS V++: Which One is Better?“ 2020 [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.bitdegree.org/tutorials/c-vs-c-plus-plus/>

[11] „Guru99“ „C++ Vs. C# – What‘s the Difference?“ [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.guru99.com/cpp-vs-c-sharp.html>

[12] „Upwork“ „C# vs. C++: Which Language is Right for Your Software Project?“ 2019 [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.upwork.com/resources/c-sharp-vs-c-plus-plus>

[13] „Educba“ „Python vs C#“ [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.educba.com/python-vs-c-sharp/>

[14] R. Shankar „C# vs Python“ [žiūrėta 07 12 2020] Prieiga per internetą: <https://hackr.io/blog/c-sharp-vs-python>

[15] „Guru99“ „Java vs C# - 10 Key Differences between Java and C#“ [žiūrėta 07 12 2020] Prieiga per internetą: https://www.guru99.com/java-vs-c-sharp-key-difference.html

[16] „Educba“ „Java vs C#“ [žiūrėta 07 12 2020] Prieiga per internetą: <https://www.educba.com/java-vs-c-sharp/>

[17] „Guru99“ „SQL vs NoSQL: What‘s the difference?“ [žiūrėta 10 12 2020] Prieiga per internetą: <https://www.guru99.com/sql-vs-nosql.html>

[18] M. Smallcombe „SQL vs NoSQL: 5 Critical Differences“ 2020 [žiūrėta 10 12 2020] Prieiga per internetą: <https://www.xplenty.com/blog/the-sql-vs-nosql-difference/>

[19] „mongoDB“ NoSQL vs SQL Databases [žiūrėta 10 12 2020] Prieiga per internetą: <https://www.mongodb.com/nosql-explained/nosql-vs-sql>

[20] M. Rouse „Microsoft SQL Server“ 2019 [žiūrėta 10 12 2020] Prieiga per internetą: <https://searchsqlserver.techtarget.com/definition/SQL-Server>

[21] „Infotec“ „What is Microsoft SQL Server and What is it Used For?“ 2017 [žiūrėta 10 12 2020] Prieiga per internetą: <https://www.infotectraining.com/blog/what-is-microsoft-sql-server-and-what-is-it-used-for>

[22] „Guru99“ „What is MongoDB? Introduction, Architecture, Features & Example” [žiūrėta 10 12 2020] Prieiga per internetą: <https://www.guru99.com/what-is-mongodb.html>

[23] „Educba“ „MongoDB vs SQL server“ [žiūrėta 10 12 2020] Prieiga per internetą: https://www.educba.com/mongodb-vs-sql-server/

[24] World wide web foundation „History of the Web“ [žiūrėta 26 12 2020] Prieiga per internetą: <https://webfoundation.org/about/vision/history-of-the-web/>

[25] D. Hemmendinger „Britannica“ „HTML“ 2000 [žiūrėta 26 12 2020] Prieiga per internetą:

<https://www.britannica.com/technology/HTML>

[26] H. S. Thompson „What‘s a URI and why does it matter?“ 2010 [žiūrėta 26 12 2020] Prieiga per internetą: <http://www.ltg.ed.ac.uk/~ht/WhatAreURIs/>

[27] R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, T. Berners-Lee „Hypertext Transfer Protocol – HTTP/1.1” 1999 [žiūrėta 26 12 2020] Prieiga per internetą: <https://www.hjp.at/doc/rfc/rfc2616.html>

[28] Md. Abu Kausar, V. S. Dhaka, Sanjeev Kumar Singh „International Journal of Computer Applications (0975 – 8887)“ 2013 „Web Crawler: A Review“ [žiūrėta 26 12 2020] Prieiga per internetą: <https://www.researchgate.net/publication/258789938_Web_Crawler_A_Review>

[29] „C#Corner“ „C# Regex Examples“ 2020[žiūrėta 09 01 2021] Prieiga per internetą: <https://www.c-sharpcorner.com/article/c-sharp-regex-examples/>

[30] „Guru99“ „What is .NET Framework? Explain Architecture & Components“ [žiūrėta 09 01 2021] Prieiga per internetą: <https://www.guru99.com/net-framework.html>

[31] „Computer Business Review“ „What is Visual Studio?“ 2021 [žiūrėta 09 01 2021] Prieiga per internetą: https://www.cbronline.com/what-is/what-is-visual-studio-4959054/#2