



VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
INFORMACINIŲ SISTEMŲ INŽINERIJOS STUDIJŲ PROGRAMA

Duomenų apie atlikėjus ir dainas rinkimo automatizavimas

Projektinio darbo ataskaita

Atliko: Lina Čerepovičiūtė, Miglė Beigaitė, Benas Sakalauskas, Veronika Mackevičiūtė

VU el. p.: lina.cerepoviciute@mif.stud.vu.lt,
migle.beigaite@mif.stud.vu.lt,
benas.sakalauskas@mif.stud.vu.lt,
veronika.mackeviciute@mif.stud.vu.lt

TURINYS

1.	ĮVADAS.....	3
1.1.	Darbo tikslas	3
1.2.	Darbo uždaviniai.....	3
1.3.	Sukurto roboto apibūdinimas	3
2.	DETALI PROCESO SCHEMA	4
2.1.	Detali esamo proceso schema.....	4
2.2.	Detali būsimo proceso schema	5
3.	ROBOTO ĮGYVENDINIMAS.....	6
3.1.	Apibendrinta programos eiga.....	6
3.2.	Atlikėjų informacijos rinkimas iš „Wikipedia“	7
3.2.1.	Atlikėjų informacijos rinkimo eiga.....	7
3.2.2.	Kiekvienam atlikėjui vykdomi veiksmai.....	7
3.3.	Informacijos rinkimas iš „YouTube“	9
3.3.1.	Informacijos rinkimo ciklo eiga	9
3.3.2.	Informacijos rinkimas faile „video.xaml“	11
3.4.	Informacijos išvedimas į „Excel“ ir grafiko formavimas	13
3.4.1.	„Excel.xaml“ failo eiga	13
3.4.2.	Veiksmai „Excel“ aplinkoje.	14
3.4.3.	Eilučių reikšmių keitimas detaliau.	14
4.	REZULTATAI	16
4.1.	„Wikipedia“ lentelė	16
4.2.	„YouTube“ lentelė	16
4.3.	Grafikas pagal prenumeratorių skaičių.	17
4.4.	Grafikas pagal patiktukų skaičių.	17
4.5.	Grafikas pagal peržiūrų skaičių.....	18
5.	GRUPĖS NARIŲ INDĖLIS.....	19
6.	APIBENDRINIMAS.....	19

1. ĮVADAS

1.1. Darbo tikslas

Automatizuoti atlikėjų ir jų kūrinių duomenų rinkimo procesą.

1.2. Darbo uždaviniai

1. Nubrėžti esamo ir būsimo proceso schemas.
2. Pagal sudarytą būsimo proceso schemą sukurti robotą UiPath Studio programoje.

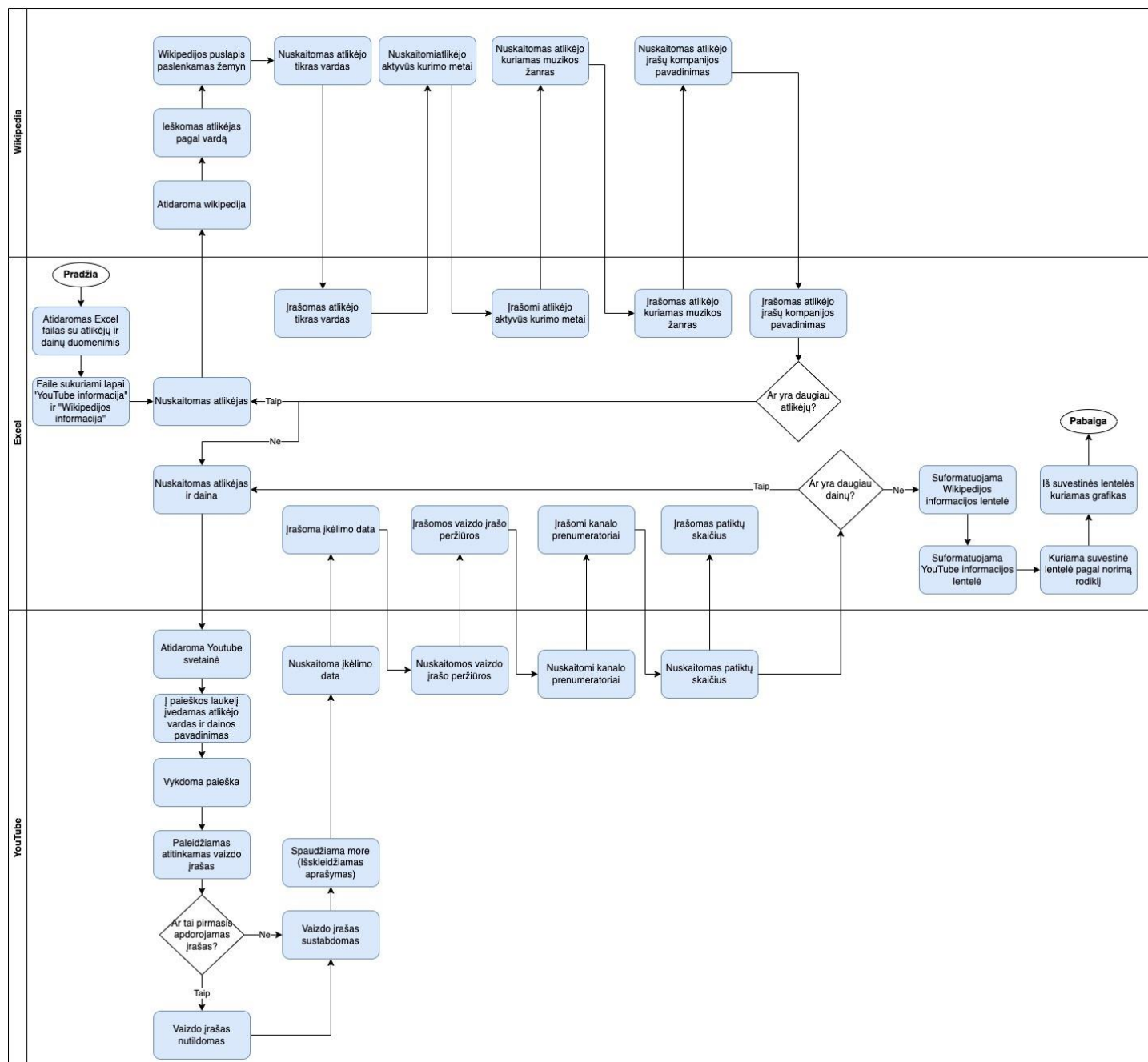
1.3. Sukurto roboto apibūdinimas

UiPath sistemoje sukurtas robotas gebantis nuskaityti pradinis duomenis (atlikėjų sceninius vardus ir kūrinių pavadinimus) iš „Excel“ failo, pagal juos rasti ir nuskaityti reikiamus duomenis iš „YouTube“ ir „Wikipedia“ tinklapių, išvesti visą surinktą informaciją į „Excel“ failą ir, pagal vartotojo pasirinktus parametrus, kitame „Excel“ dokumento puslapyje sukurti grafiką.

2. DETALI PROCESO SCHEMA

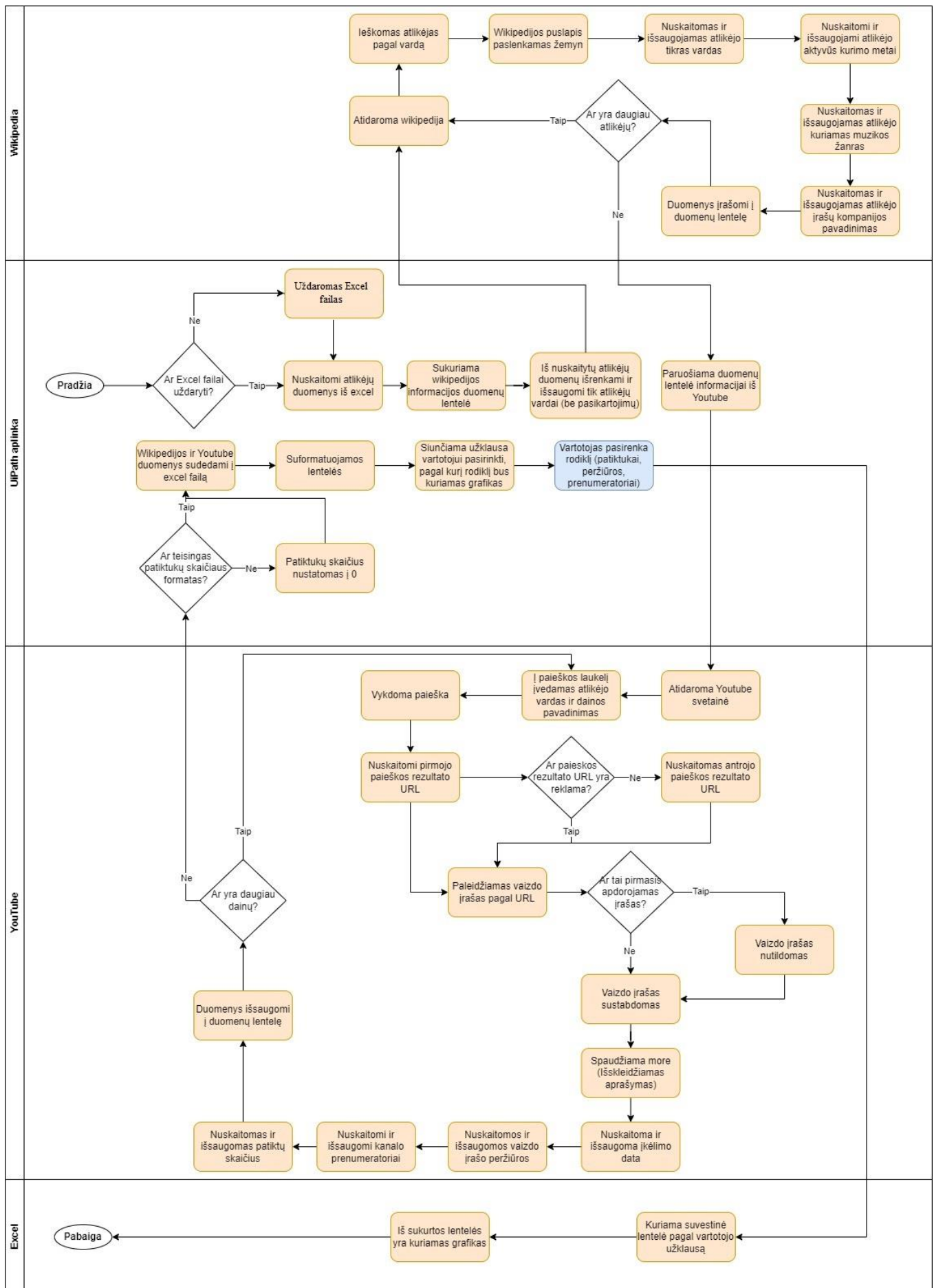
Oranžine spalva žymimi automatizuoti žingsniai, mėlyna – žmogaus atliekami veiksmai.

2.1. Detali esamo proceso schema



Pav. 1 Detali esamo proceso schema

2.2. Detali būsimo proceso schema

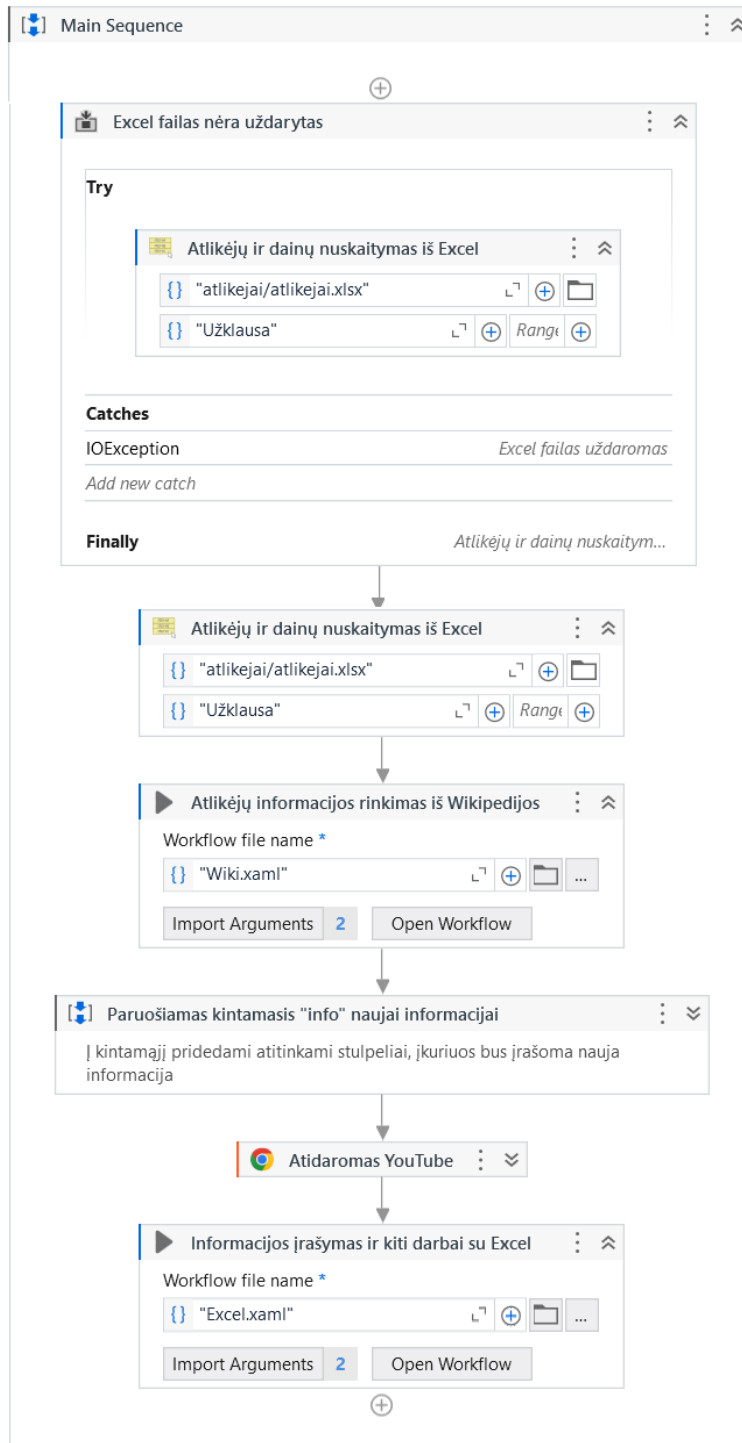


Pav. 2 Detali būsimo proceso schema

3. ROBOTO ĮGYVENDINIMAS

3.1. Apibendrinta programos eiga

Programos pradžioje patikrinama ar excel failas iš kurio bus nuskaitymi duomenys nėra atidarytas. Jei failas atidarytas – jis uždarytas. Tuomet iš failo nuskaitymi pradiniai duomenys (atlikėjų vardai ir dainų pavadinimai). Programa surenka reikiamus duomenis iš „Wikipedia“ puslapio. Toliau, kintamasis „info“ paruošiamas duomenų rinkimui iš „YouTube“ puslapio – pridedami stulpeliai „Įkėlimo data“, „Peržiūros“, „Prenumeratoriai“ ir „Patiktukų skaičius“. Surinkus reikiamus duomenis iš „YouTube“ puslapio, jie įrašomi į „Excel“ failą ir, pagal vartotojo pasirinkimus, tame pačiame faile sukuriamas grafikas.



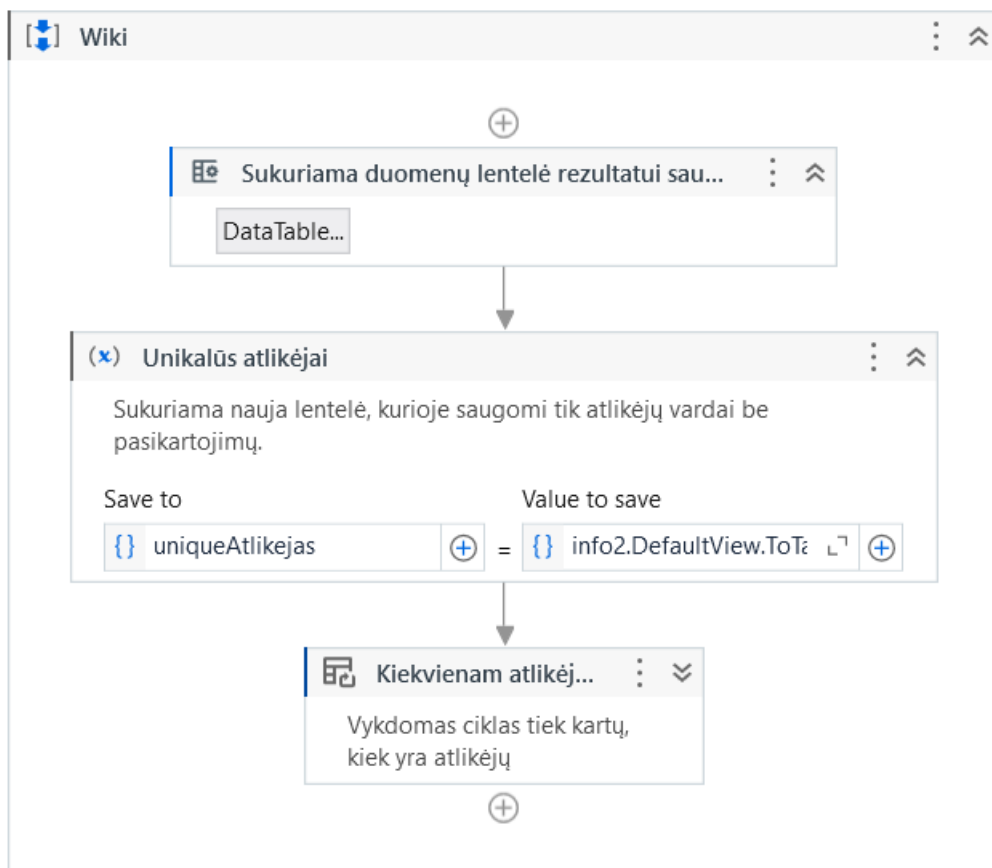
Pav. 3 Apibendrinta programos eiga

3.2. Atlikėjų informacijos rinkimas iš „Wikipedia“

Atlikėjų informacijos rinkimas vykdomas atskirame „Workflow“ faile „Wiki.xaml“.

3.2.1. Atlikėjų informacijos rinkimo eiga

Pirmiausia sukuriami duomenų lentelė atlikėjų duomenų saugojimui. Ją sudaro 5 stulpeliai – „Atlikėjas“, „Atlikėjo tikras vardas“, „Kuriamos muzikos žanras“, „Aktyvūs metai“ ir „Įrašų kompanijos pavadinimas“. Tuomet, sukuriami nauja lentelė, kurioje saugomi tik unikalūs įrašai. Galiausiai, surenkama reikiama informacija apie kiekvieną atlikėją.



Pav. 4 Atlikėjų informacijos rinkimo eiga

3.2.2. Kiekvienam atlikėjui vykdomi veiksmai

Kiekvieno atlikėjo informacijos surinkimui atliekami tokie veiksmai:

1. Į paieškos laukelį įvedamas atlikėjo vardas.
2. Spaudžiamas paieškos mygtukas.
3. Puslapis paslenkamas žemyn.
4. Nuskaitomas atlikėjo tikrasis vardas.
5. Nuskaitomi atlikėjo aktyvūs metai.
6. Nuskaitomas atlikėjo muzikos žanras.
7. Nuskaitomas atlikėjo įrašų kompanijos pavadinimas.
8. Sukuriama duomenų eilutė su nuskaitytais duomenimis.
9. Duomenų eilutė pridedama į lentelę.



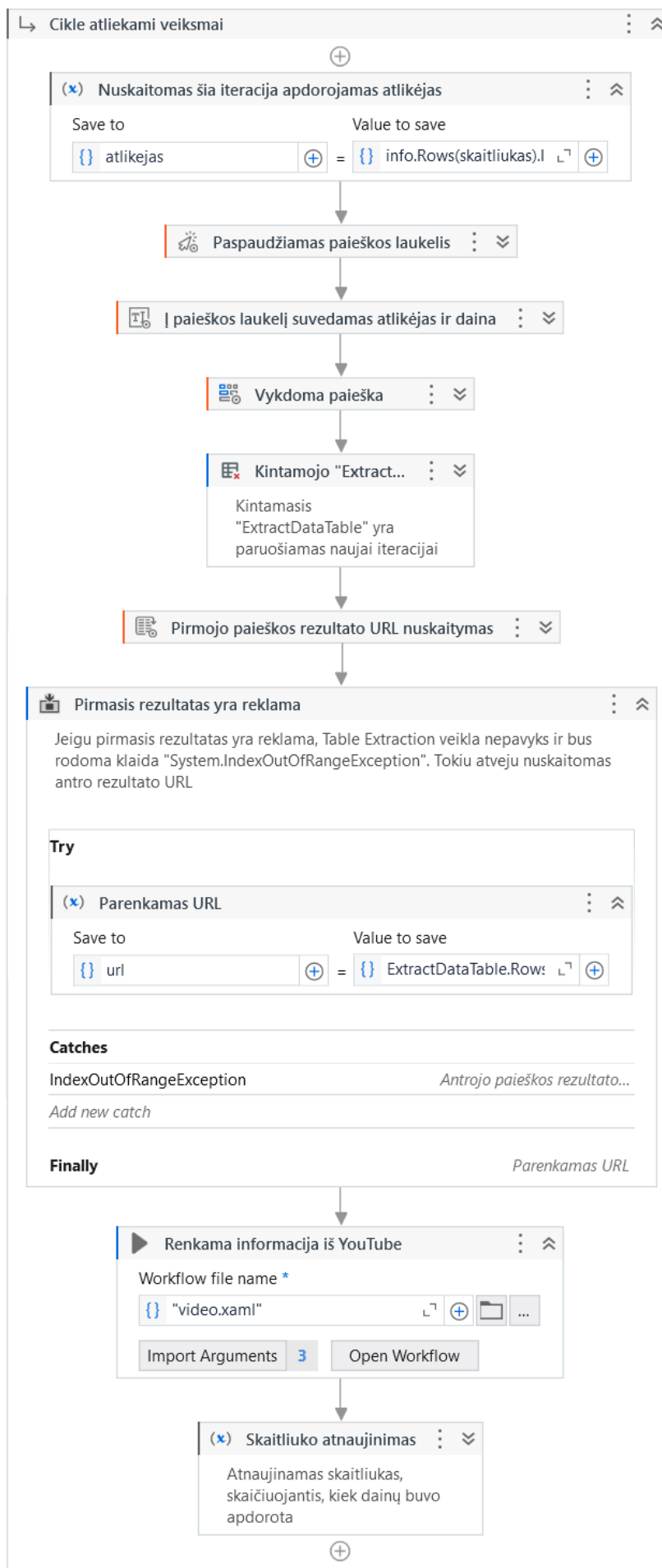
Pav. 5 Kiekvienam atlikėjui atliekami veiksmai

3.3. Informacijos rinkimas iš „YouTube“

3.3.1. Informacijos rinkimo ciklo eiga

Informacijos rinkimui iš „YouTube“ tinklapio cikle vykdomi tokie veiksmai:

1. Nuskaitomas vardas atlikėjo, kurio informacija pildoma.
2. Paspaudžiamas paieškos laukelis.
3. Į paieškos laukelį įvedamas atlikėjo vardas ir daina.
4. Paspaudžiamas „Enter“ klavišas.
5. Išvalomas kintamasis „ExtractDataTable“.
6. Į „ExtractDataTable“ nuskaitoma pirmojo paieškos rezultato nuoroda.
7. Patikrinama ar pirmasis rezultatas nėra reklama. (Jei reklama – nuskaitoma antrojo paieškos rezultato nuoroda.)
8. Vykdomas failas „video.xaml“.
9. Vienetu padidinamas ciklo skaitliukas.

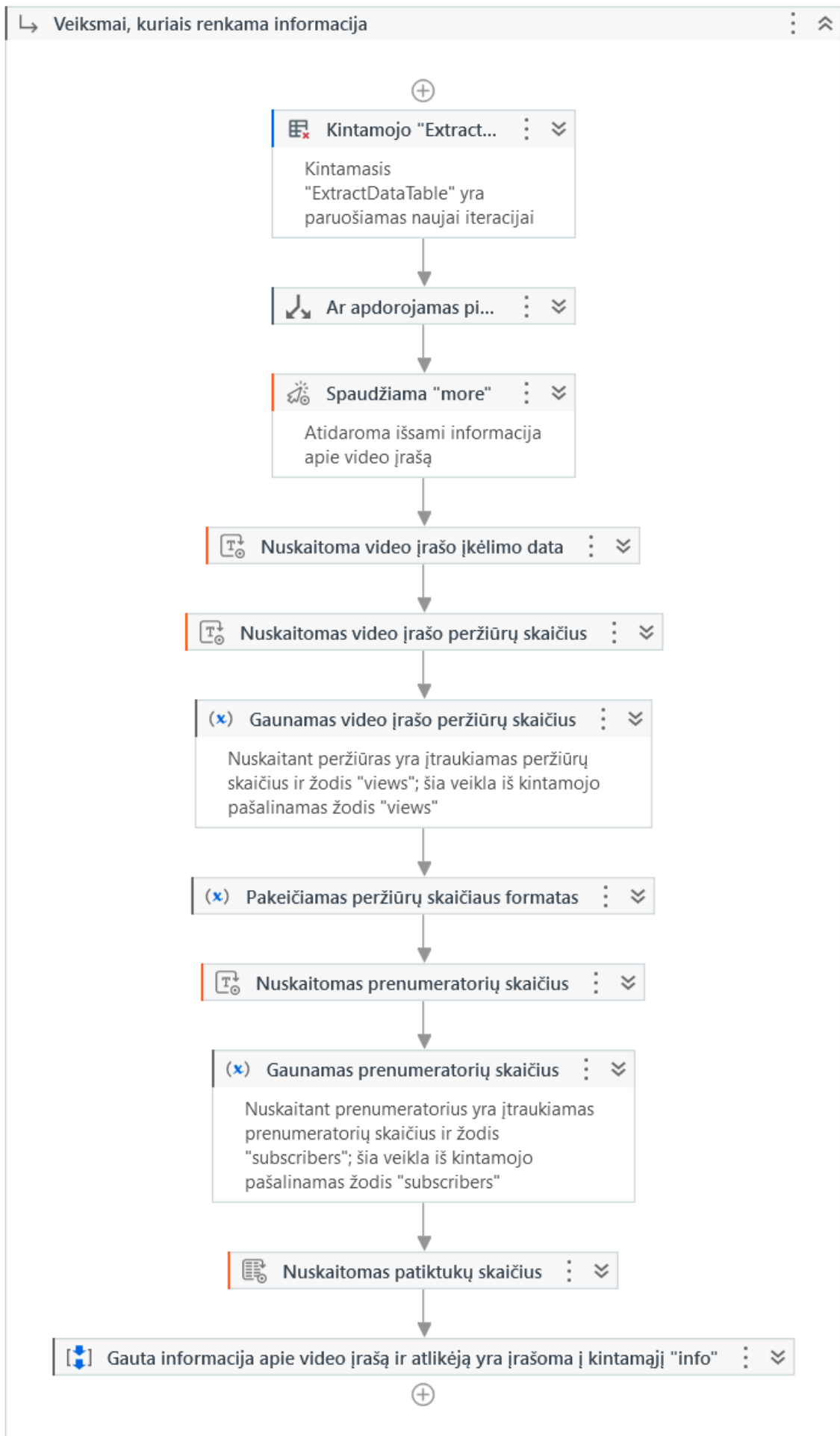


Pav. 6 Informacijos rinkimo iš "YouTube" ciklas

3.3.2. Informacijos rinkimas faile „video.xaml“

Failo „video.xaml“ pradžioje paleidžiamas „YouTube“ vaizdo įrašas. Toliau vykdomi informacijos rinkimo veiksmai:

1. Išvalomas kintamasis „ExtractDataTable“.
2. Patikrinama ar apdorojamas pirmasis vaizdo įrašas:
 - 2.1. Jei apdorojamas pirmasis vaizdo įrašas – spaudžiamas klavišas „M“ (vaizdo įrašo garso išjungimui) ir „K“ (vaizdo įrašo sustabdymui).
 - 2.2. Jei apdorojamas vaizdo įrašas nėra pirmasis – spaudžiamas tik klavišas „K“.
3. Spaudžiamas mygtukas „more“.
4. Nuskaitoma vaizdo įrašo įkėlimo data.
5. Nuskaitomas vaizdo įrašo peržiūrų skaičius.
6. Iš nuskaityto teksto pašalinamas žodis „views“.
7. Peržiūrų skaičiaus formatas pakeičiamas iš „string“ į „int“.
8. Nuskaitomas prenumeratorių skaičius.
9. Iš nuskaityto teksto pašalinamas žodis „subscribers“.
10. Nuskaitomas patiktukų skaičius.
11. Surinkta informacija įrašoma į kintamąjį „info“.



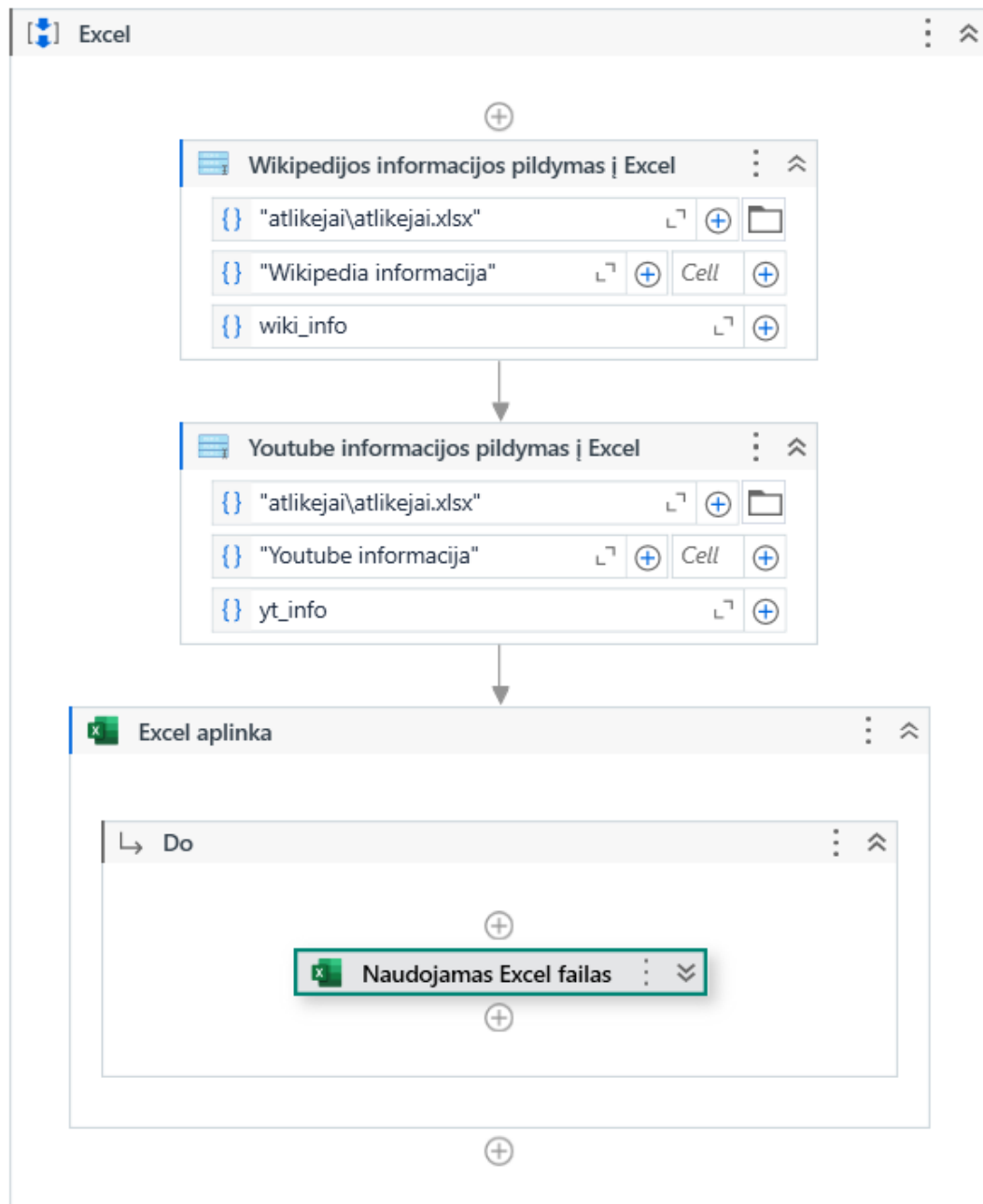
Pav. 7 Informacijos rinkimas

3.4. Informacijos išvedimas į „Excel“ ir grafiko formavimas

Informacijos išvedimas į „Excel“ failą ir grafiko kūrimas vykdomas faile „Excel.xaml“.

3.4.1. „Excel.xaml“ failo eiga

Pirmiausia į „Excel“ failą įrašoma informacija surinkta iš „Wikipedia“ tinklapio. Tuomet į tą patį failą įrašoma informacija surinkta iš „YouTube“ tinklapio. Galiausiai atliekami veiksmai „Excel“ aplinkoje.

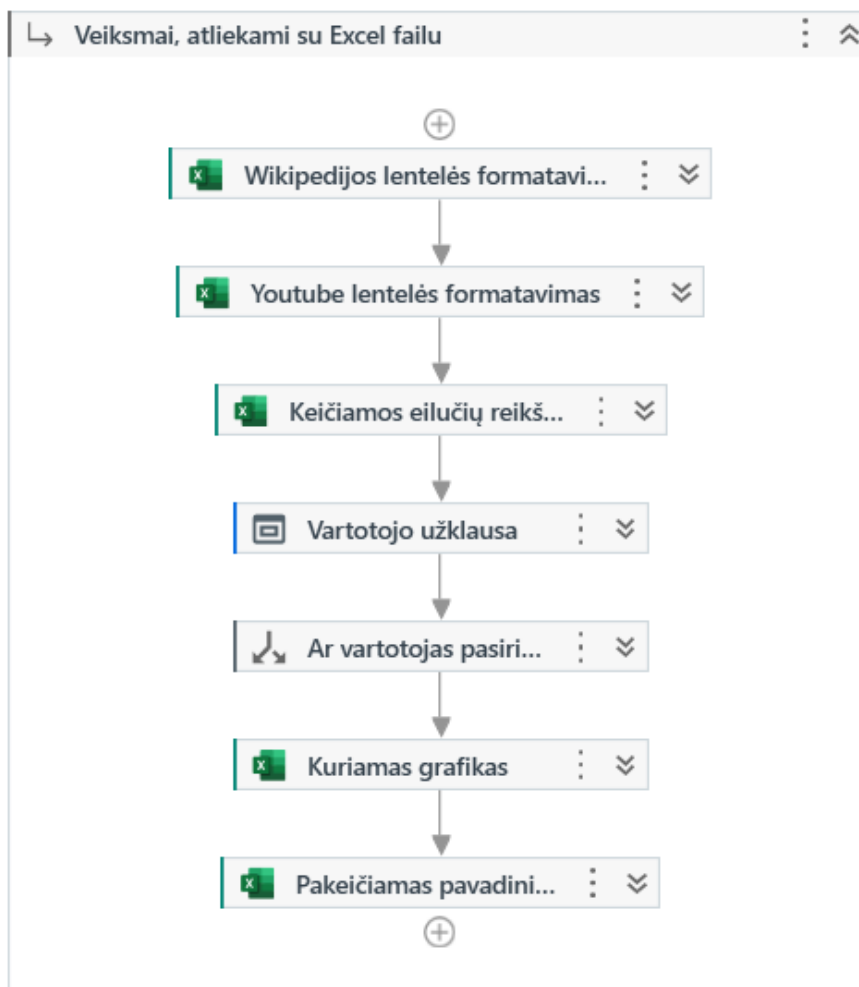


Pav. 8 Informacijos išvedimas į "Excel"

3.4.2. Veiksmai „Excel“ aplinkoje.

„Excel“ aplinkoje atliekami tokie veiksmai:

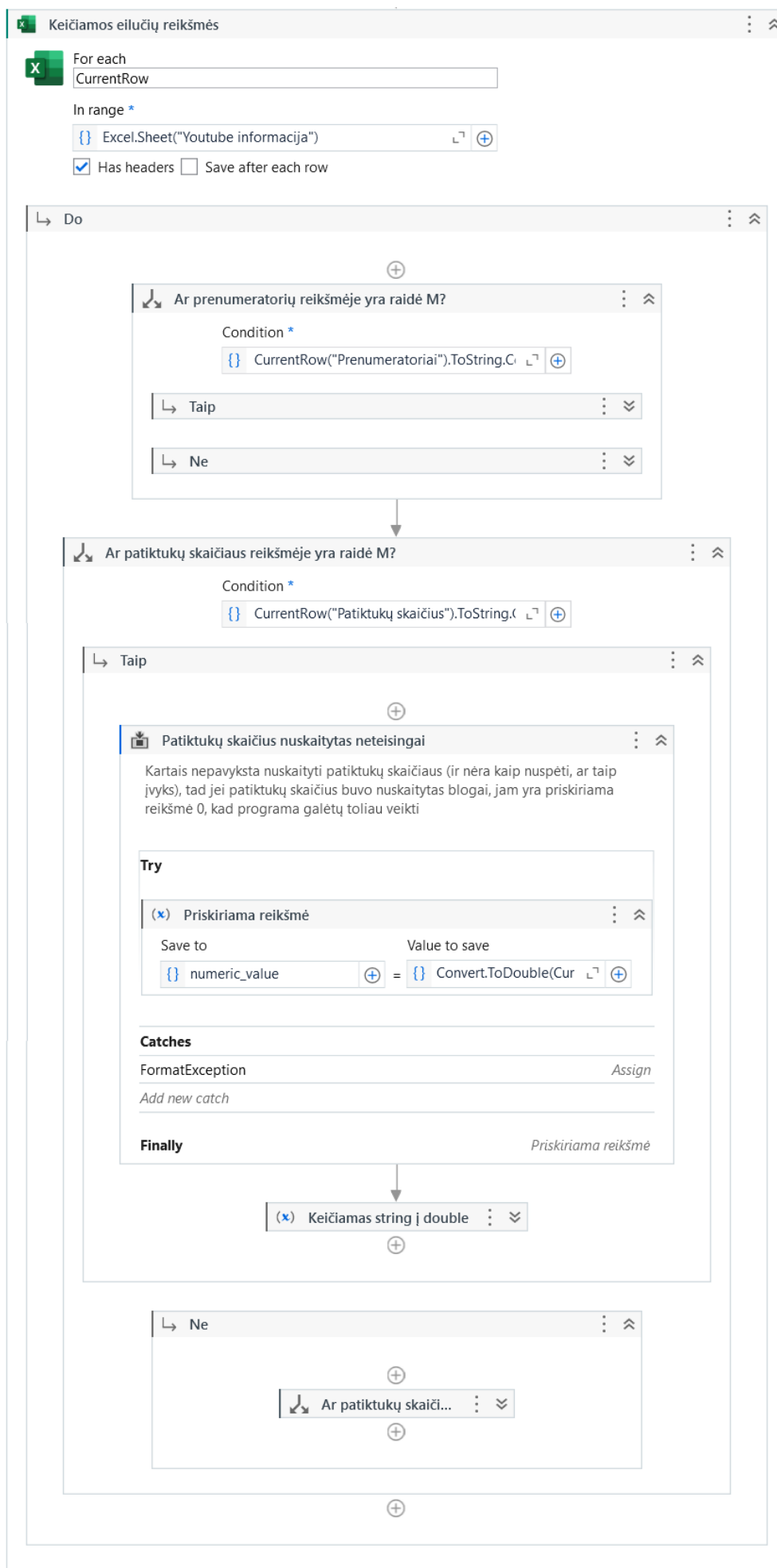
1. Suformatuojama „Wikipedia“ informacijos lentelė.
2. Suformatuojama „YouTube“ informacijos lentelė.
3. Į skaičius pakeičiamos prenumeratorių laukelio reikšmės. („M“ = 1 000 000, „K“ = 1000)
4. Vartotojui pateikiama grafiko pasirinkimo užklausa.
5. Pagal vartotojo pasirinkimą sukuriamas atitinkama suvestinė lentelė. (Prenumeratorių, peržiūrų arba patiktukų skaičiaus)
6. Sukuriamas grafikas.
7. Pakeičiamas grafiko pavadinimas.



Pav. 9 Veiksmai "Excel" aplinkoje

3.4.3. Eilučių reikšmių keitimas detaliau.

Pirma pakeičiamos prenumeratorių laukelio reikšmės (vietoj „M“ ir „K“ raidžių reikšmės pakeičiamos skaičiais). Tuomet bandoma pakeisti patiktukų laukelio reikšmės. Jei nepavyksta tinkamai pakeisti reikšmių į skaičius, įrašoma reikšmė 0, kad programa galėtų veikti toliau.



Pav. 10 Eilučių reikšmių keitimo žingsniai

4. REZULTATAI

Sukurtam „UiPath“ robotui surinkus ir apdorojus „Excel“ faile pateiktą atlikėjų informaciją gauti tokie rezultatai:

4.1. „Wikipedia“ lentelė

	A	B	C	D	E
1	Atlikėjas	Atlikėjo tikras vardas	Kuriamos muzikos žanras	Aktyvūs metai	Jrašų kompanijos pavadinimas
2	Kendrick Lamar	Kendrick Lamar Duckworth	West Coast hip hop	2003–present	Top Dawg
3	Lady Gaga	Stefani Joanne Angelina Germanotta	Pop	2001–present	Def Jam
4	Cardi B	Belcalis Marlenis Almánzar	Hip hop	2014–present	Hip hop
5	Rihanna	Robyn Rihanna Fenty	R&B	2003–present	Roc Nation

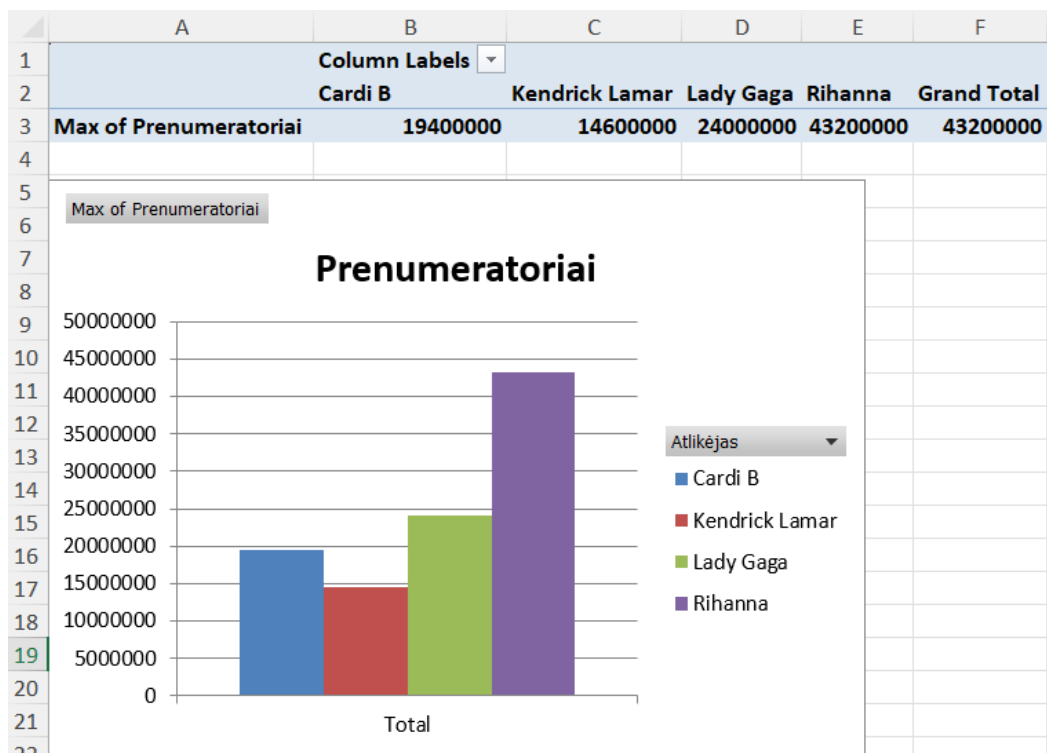
Pav. 11 Informacija surinkta iš puslapio "Wikipedia"

4.2. „YouTube“ lentelė

	A	B	C	D	E	F
1	Atlikėjas	Dainos pavadinimas	Įkėlimo data	Peržiūros	Prenumeratoriai	Patiktųjų skaičius
2	Kendrick Lamar	King Kunta	Apr 2, 2015	174704154	14600000	1100000
3	Lady Gaga	Americano	May 19, 2011	33601669	24000000	300000
4	Cardi B	Bodak Yellow	Jun 25, 2017	1.136E+09	19400000	8200000
5	Lady Gaga	Just Dance	Jun 17, 2009	444623147	24000000	2000000
6	Kendrick Lamar	Swimming Pools	Aug 3, 2013	200829316	14600000	1900000
7	Lady Gaga	Alejandro	Jun 8, 2010	556875505	24000000	2600000
8	Rihanna	Don't Stop The Music	Nov 21, 2009	708820511	43200000	2500000
9	Cardi B	Be Careful	May 21, 2018	174857196	19400000	1700000
10	Kendrick Lamar	Rich Spirit	Nov 16, 2022	32769575	14600000	677000
11	Lady Gaga	Heavy Metal Lover	May 19, 2011	18974045	24000000	189000
12	Rihanna	SOS	Oct 7, 2009	142359282	43200000	525000
13	Rihanna	We Found Love	Oct 20, 2011	1.123E+09	43200000	5000000
14	Cardi B	I Like It	May 29, 2018	1.618E+09	19400000	10000000
15	Kendrick Lamar	N95	May 14, 2022	81198646	14600000	1600000
16	Rihanna	Only Girl	Oct 13, 2010	947927989	43200000	3300000
17	Cardi B	Ring	Aug 20, 2018	249783517	19400000	2100000

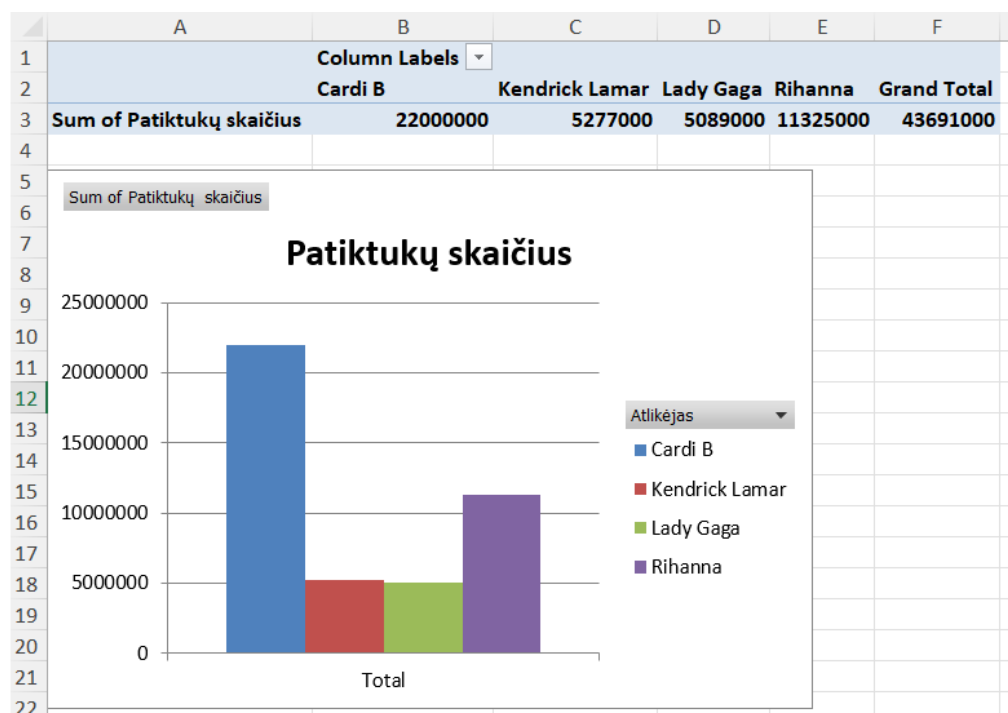
Pav. 12 Informacija surinkta iš "YouTube"

4.3. Grafikas pagal prenumeratorių skaičių.



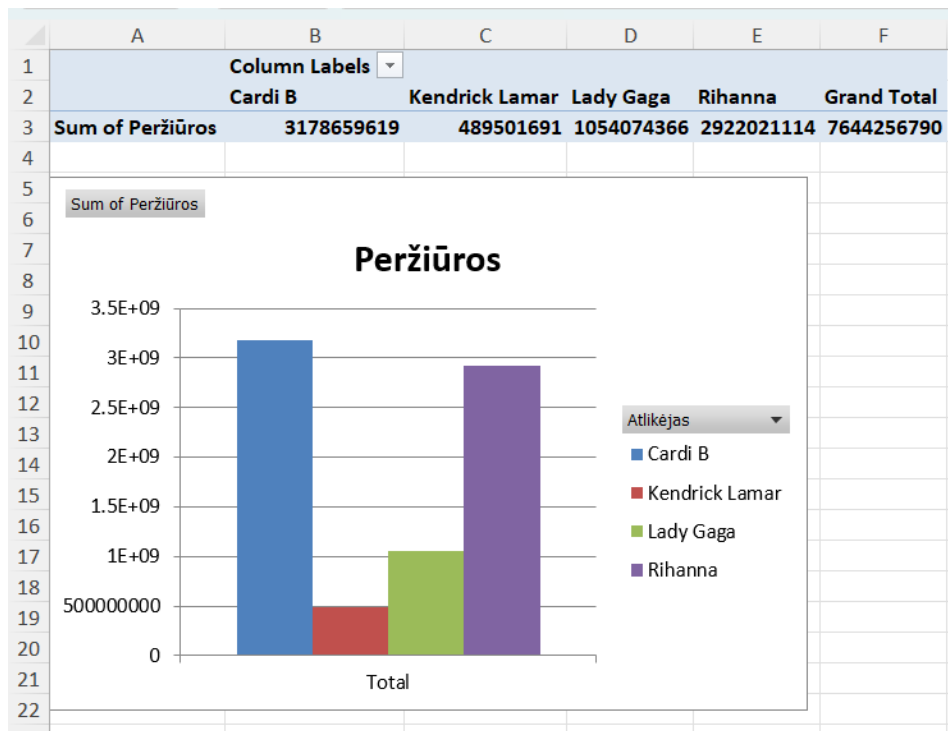
Pav. 13 Prenumeratorių grafikas

4.4. Grafikas pagal patiktukų skaičių.



Pav. 14 Patiktukų grafikas

4.5. Grafikas pagal peržiūrų skaičių.



Pav. 15 Peržiūrų grafikas

5. GRUPĖS NARIŲ INDĖLIS

1. Veronika Mackevičiūtė – Informacijos rinkimo iš „Wikipedia“ automatizavimas.
2. Miglė Beigaitė – Informacijos nuskaitymo iš „Excel“ failo ir informacijos rinkimo iš „YouTube“ automatizavimas.
3. Benas Sakalauskas – Informacijos išvedimo į „Excel“ failą ir grafiko kūrimo automatizavimas. Proceso schemų kūrimas.
4. Lina Čerepovičiūtė – Ataskaitos rengimas.

6. APIBENDRINIMAS

Siekiant darbo tikslo automatizuoti atlikėjų ir jų kūrinių duomenų rinkimo procesą, buvo iškelti ir atlikti du darbo uždaviniai:

1. Nubrėžtos esamo ir būsimą proceso schemas.
2. Pagal sudarytą būsimą proceso schemą sukurtas robotas UiPath Studio programoje.

Realizuojant robotą buvo išspręstos visos iškilusios problemos, pasiektas planuotas funkcionalumas.