әл-Фараби атындағы Қазақ ұлттық университеті
География және табиғатты пайдалану факультеті
Рекреациялық география және туризм кафедрасы

9-дәріс

Туристердің цифрлық іздері және оларды жинау технологиялары

Қонақ үй және туристік бизнестегі ақпараттық технологиялар

ДӘРІСКЕР

Елдар Нұрұлы

Аға оқытушы, аға ғылыми қызметкер, PhD Candidate



Анталиядағы бір күн: қанша із?

- **08:10** Google-да «Kaleiçi museum» іздеу
- **09:30** телефон базалық станцияға қосылды
- **12:15** кафеде картамен төлем
- **15:40** геотегі бар фото Instagram-да
- **21:00** Google Maps-та мейрамханаға пікір



Фото: DimiTalen, CC0, Wikimedia Commons

Сұрақ: осы іздерді кім жинайды және туризм бизнесі оларды қалай пайдаланады?

Дәріс жоспары

01 **Цифрлық із деген не?**
белсенді және енжар іздер, деректер түрлері

02 **Жинау технологиялары**
API, веб-скрейпинг, мобильді деректер, іздеу сұраныстары

03 **Шекаралар**
заң, платформа ережелері, деректер сапасы

04 **Қазақстан тәжірибесі**
мобильді деректер негізіндегі туризм статистикасы

05 **Семинарға көпір**
Instant Data Scraper-мен жұмыс жоспары

Негізгі сұрақ

Туристің цифрлық ізі — қонақ үй үшін құнды ақпарат па, әлде қорғауды қажет ететін жеке дерек пе?

Дәрістен кейін сіз...



Цифрлық іздің түрлерін ажырата аласыз: белсенді/енжар, сапарға дейін/кезінде/кейін



Әр міндетке сай жинау технологиясын таңдай аласыз



Деректерді жинаудың заңдық және этикалық шекарасын түсіндіре аласыз



Семинарда өз нысаныңыз туралы деректер жинауға дайын боласыз

01

Цифрлық із деген не?

Белсенді және енжар іздер, сапар кезеңдері, деректер түрлері



Цифрлық із: белсенді және енжар

Цифрлық із — адам интернетте, құрылғыларда және цифрлық қызметтерде қалдыратын деректер: орны, уақыты, әрекеті, пікірі.

Белсенді іздер

турист өзі саналы түрде жариялайды

- пікірлер мен рейтингтер
- геотегі бар фото мен видео
- блог, әлеуметтік желі жазбалары

Енжар іздер

жүйе туристке білдірмей тіркейді

- ұялы байланыс операторының деректері
- Wi-Fi және Bluetooth журналдары
- карта төлемдері, сайтқа кіру журналы

Дереккөз: Girardin et al., 2008 [1]

Сапарға дейін, сапар кезінде, сапардан кейін



Сапарға дейін

- іздеу сұраныстары
- карталар мен сайттарды қарау
- брондау



Сапар кезінде

- ұялы желіге қосылу
- карта төлемдері
- геотегі бар фото



Сапардан кейін

- пікірлер мен бағалар
- фотоальбомдар
- блог, видео



Girardin және әріптестері (2008) Римдегі туристерді Flickr фотолары мен ұялы байланыс деректері арқылы «көрді»

Дереккөз: Girardin et al., 2008 [1]

Туризмдегі үлкен деректердің үш көзі



Пайдаланушы жасаған контент (UGC)

мәтін: пікірлер, блогтар · фото: геотегі бар суреттер

Мысалы: Tripadvisor, Google Maps, Instagram



Құрылғы деректері

GPS, мобильді роуминг, Bluetooth, Wi-Fi, RFID

Мысалы: ұялы операторлар, датчиктер



Операция (транзакция) деректері

іздеу сұраныстары, сайтқа кіру, онлайн брондау

Мысалы: Google Trends, Booking, қонақ үй сайты

Дереккөз: Li et al., 2018 [2]

Пікірлер көп — бірақ бәрі шынайы ма?

31,1 млн

пікір Tripadvisor-ге 2024 жылы жіберілді

2,7 млн

жалған пікір бұғатталды

54%

алаяқтықтың үлесі — рейтингті жасанды көтеру («review boosting»)

Қорытынды: цифрлық із — бай дерек көзі, бірақ «шикізат». Оны жинағаннан кейін тексеру, тазалау және сыни бағалау қажет — семинардың мақсаты да осы

Дереккөз: Tripadvisor, 2025 [3]

Белсенді ме, енжар ма?

Белсенді із — оң қол, енжар із — сол қол

1. Tripadvisor-ге жазылған пікір
2. Ұялы оператор тіркеген базалық станция
3. Instagram-дағы геотегі бар фото
4. Қонақ үй Wi-Fi желісіне қосылу журналы
5. Кафеде банк картасымен төлеу
6. TikTok-тағы саяхат туралы видео

Белсенді ме, енжар ма: жауаптар

Tripadvisor-ге жазылған пікір	**Белсенді**	турист өзі жариялады
Ұялы оператор тіркеген базалық станция	**Енжар**	телефон желіге автоматты түрде қосылды
Instagram-дағы геотегі бар фото	**Белсенді**	орнын турист өзі белгіледі
Қонақ үй Wi-Fi желісіне қосылу журналы	**Енжар**	жүйе өздігінен тіркейді
Кафеде банк картасымен төлеу	**Енжар**	операцияның «жанама» ізі
TikTok-тағы саяхат туралы видео	**Белсенді**	контентті турист жасады

Дереккөз: Girardin et al., 2008 [1]

02

Цифрлық іздерді қалай жинайды?

API, веб-скрейпинг, мобильді деректер, іздеу сұраныстары, геотегті фото



Бес негізгі технология



API

жүйелер арасындағы ресми «келісім»



Веб-скрейпинг

ашық веб-беттерден кестеге жинау



Мобильді деректер

ұялы операторлардың иесіздендірілген деректері



Іздеу сұраныстары

Google Trends: адамдар не іздейді



Геотегті контент

фото мен видеоның орны мен уақыты

Семинарда — веб-скрейпинг

Дереккөз: Li et al., 2018 [2]

API мен веб-скрейпинг: айырмашылығы неде?

	API	Веб-скрейпинг
Кім береді?	платформа өзі, ереже бойынша	беттің көрінетін бөлігінен аламыз
Тұрақтылық	жоғары: формат өзгермейді	төмен: сайт дизайны өзгерсе, істен шығады
Рұқсат	келісім мен кілт (key) керек	сайт ережелері мен заңды тексеру керек
Құны	жиі ақылы немесе шектеулі	құралдар көбіне тегін
Мысал	Booking → қонақ үй жүйесі (Channel manager)	Instant Data Scraper → Excel

Веб-скрейпинг қалай жұмыс істейді?



Бетті ашу

браузер немесе бағдарлама сайтқа сұраныс жібереді



HTML-ді оқу

беттің құрылымы: тақырып, баға, пікір



Қажеттісін бөліп алу

қайталанатын блоктарды табу



Кестеге сақтау

CSV / XLSX файл

Instant Data Scraper — Chrome кеңейтімі: бетте кестені өзі табады («Try another table»), келесі беттерге өтеді («Locate "Next" button» → «Start crawling»), XLSX/CSV форматында жүктеп алады. Барлығы сіздің браузеріңізде жүреді.

Дереккөз: Byteful, 2026; Geonode, 2026 [4]

Ұялы деректер туристі қалай «көреді»?



Иесіздендірілген ID
аты-жөні жоқ, кездейсоқ нөмір

Уақыт пен базалық станция
қай жерде, қашан

SIM-картаның ел коды
шетелдік турист пе?

«Әдеттегі орта»
соңғы 12 айдағы мекеннен тыс шығу = сапар

Эстония бұл әдісті туризм статистикасында ертеден қолданады; Қазақстан 2026 жылдан бастап енгізуде

Дереккөздер: Economist.kg, 2026 [5]; Data Collaboratives [6]

Іздеу сұраныстары мен фотолар не айтады?

Google Trends → сұраныс болжамы

≈ 30%

Багам аралдарына АҚШ туристерінің келуін болжаудағы қате Google Trends қосылғанда осыншаға азайды (ХВҚ, 2020)

Геотегті фото → туристер «картасы»

Рим: Flickr-дегі ашық фотолар мен ұялы желі деректерін біріктіріп, туристердің қай жерде шоғырланатыны мен қалай қозғалатыны көрсетілді (Girardin et al., 2008)

Дереккөздер: IMF WP/20/22, 2020 [7]; Girardin et al., 2008 [1]

Қай технология көмектеседі?

Нұсқалар: веб-скрейпинг · API · мобильді деректер · іздеу сұраныстары · геотегті фото

1. Анталиядағы 50 қонақ үйдің рейтингін кестеге жинау
2. Шетелдік туристер санын аймақтар бойынша бағалау
3. Келесі маусымдағы сұранысты болжау
4. Booking-тегі брондауды қонақ үй жүйесіне автоматты түрде жіберу
5. Қаладағы ең танымал суретке түсетін танымал орындарды табу

Қай технология: жауаптар

Анталиядағы 50 қонақ үйдің рейтингін кестеге жинау	**Веб-скрейпинг**	ашық беттен кестеге
Шетелдік туристер санын аймақтар бойынша бағалау	**Мобильді деректер**	иесіздендірілген SIM-деректер
Келесі маусымдағы сұранысты болжау	**Іздеу сұраныстары**	Google Trends
Booking-тегі брондауды қонақ үй жүйесіне автоматты түрде жіберу	**API**	жүйелер ресми «келісім» бойынша деректер алмасады
Қаладағы ең танымал суретке түсетін танымал орындарды табу	**Геотегті фото**	UGC: Flickr, Instagram

Дереккөздер: Li et al., 2018 [2]; IMF WP/20/22, 2020 [7]

03

Шекаралар: заң, ереже, сапа

«Техникалық мүмкін» әрқашан «рұқсат етілген» дегенді білдірмейді



ҚР дербес деректер заңы: 2026 жылғы өзгерістер

Жариялау ≠ келісім	25.08.2026 ж. бастап: адам деректерін өзі жариялаған болса да, бұл оларды кейін жинауға автоматты келісім емес
Бүркемелеу (маскалау)	ашық таратуға болмайтын мәліметтер бүркемеленуі тиіс
Тек автоматты шешімге тыйым	тек автоматты өңдеуге негізделген шешімдерге тыйым; адамның қарсылық білдіру құқығы бар

Семинар ережесі: пікір авторларының аты-жөнін, профиль сілтемелерін, фотоларын жинамаймыз

Дереккөз: «Дербес деректер туралы» Заң; Morgan Lewis, 2026 [8]

Сайттың да өз ережесі бар

Google Maps шарттары (2-бөлім)

тыйым салынады:

- Google Maps бөліктерін қайта таратуға, сатуға
- контентті рұқсатсыз көшіруге
- контентті жаппай жүктеуге (mass download)

Оқу жобасында не істейміз?

- шағын іріктеме: 30–100 жол
- тек оқу мақсатында, жарияламаймыз
- жеке деректерді бірден жоямыз
- сайттың ережесі мен robots.txt файлын оқып, есепте көрсетеміз

Дереккөз: Google Maps Terms, 2026 [9]

Цифрлық із кімнің дауысы — және кімдікі емес?

Жалған пікірлер

Tripadvisor 2024 жылы 2,7 млн жалған пікірді бұғаттады

Іріктеме ығысуы

пікірді бәрі жазбайды: көбіне өте риза немесе өте наразы адамдар

Тіл мен платформа

түрік, орыс, неміс тіліндегі пікірлер әр платформада әртүрлі

Цифрлық теңсіздік

смартфоны жоқ немесе әлеуметтік желі қолданбайтын туристер «көрінбейді»

Дереккөз: Tripadvisor, 2025 [3]

Деректерді жинамас бұрын: 5 сұрақ

Бұл заңды ма?	дербес деректер туралы заң
Сайт ережесі не дейді?	Terms of Use, robots.txt
Жеке дерек бар ма?	аты-жөн, фото, профиль — жоямыз
Минимум жинап отырмын ба?	тек мақсатқа қажетті бағандар
Қайда сақтаймын?	жарияламаймын, оқудан кейін жоямын

Шын ба, жалған ба?

Шын — оң қол, жалған — сол қол

1. Пікір ашық тұрса, оны кез келген мақсатта жинап, тарата беруге болады
2. Техникалық тұрғыдан жинауға болатынның бәрі заңды
3. Tripadvisor 2024 жылы 2,7 млн жалған пікірді бұғаттады
4. Ұялы оператор деректерінде туристің аты-жөні болады
5. Google іздеу деректері туристер ағынын болжауды жақсарта алады

Шын ба, жалған ба: жауаптар

Пікір ашық тұрса, оны кез келген мақсатта жинап, тарата беруге болады	**ЖАЛҒАН**	ҚР заңы (2026): жариялау — жинауға келісім емес; платформа ережелері де шектейді
Техникалық тұрғыдан жинауға болатынның бәрі заңды	**ЖАЛҒАН**	«мүмкін» ≠ «рұқсат етілген»
Tripadvisor 2024 жылы 2,7 млн жалған пікірді бұғаттады	**ШЫН**	компанияның 2025 жылғы есебі
Ұялы оператор деректерінде туристің аты-жөні болады	**ЖАЛҒАН**	тек иесіздендірілген идентификатор, уақыт, станция, ел коды
Google іздеу деректері туристер ағынын болжауды жақсарта алады	**ШЫН**	Багам аралдары: болжам қатесі ≈ 30%-ға азайды (ХВҚ)

Дереккөздер: «Дербес деректер туралы» Заң; Morgan Lewis, 2026 [8]; Tripadvisor, 2025 [3]; Economist.kg, 2026 [5]; IMF WP/20/22, 2020 [7]

04

Қазақстан: цифрлық іздер статистикада

Туристерді мобильді деректер арқылы есептеу



Туристерді мобильді деректер бойынша санау

15,3 млн

шетелдік келуші 2024 жылы

10,4 млн

олардың бір тәуліктен көп болғаны

≈ 9 млн

орналастыру орындарында тіркелмеген

Шешім

2026 жылдан бастап ішкі және келу туризмін иесіздендірілген мобильді деректер негізінде есептеу. Дәстүрлі статистика сақталады, жаңа дерек оны толықтырады.

Сыни сұрақ

Пилоттық жобада бір ғана оператордың деректері қолданылды. Бұл барлық туристі көрсете ала ма?

Дереккөздер: Kazinform [10]; Economist.kg, 2026 [5]

Сіздің қонақ үйіңіздің цифрлық іздері қайда?



Google Maps

рейтинг, пікірлер, фото



Tripadvisor

пікірлер, сапар түрі



Booking.com

рейтинг, баға, пікір



Instagram, TikTok

геотег, видео



Қонақ үй сайты

нөмір түрлері, баға

Тапсырма: Түркиядағы қонақ үйіңізді немесе Қазақстандағы нысанды таңдап, оның іздері бар 3 платформаны тауып келіңіз — семинарда солардың бірінен деректер жинаймыз

Семинар 9: не істейміз?

1	**Орнату**	Chrome-ға Instant Data Scraper
2	**Нысан**	Түркиядағы немесе Қазақстандағы қонақ үй
3	**Жинау**	кемінде 30 пікір (Google Maps) → XLSX
4	**Тазалау**	жеке деректі жою, бағандар, баға, күн, қайталанған жолдар
5	**Алғашқы талдау**	орташа баға, бағалардың таралуы



Фото: Matti Blume, CC BY-SA 4.0, Wikimedia Commons

Мысалы: Шымбұлақ қонақ үйі

Нұсқаулық пен видеолар — семинар тапсырмасында

ВИДЕО

Қарап шығыңыз: 3 қысқа видео

Веб-скрейпинг деген не?
compas pro IT · орыс тілі

Instant Data Scraper: қадамдар
Tim Yakubson · ағылшын тілі

Google Maps тізімін жинау
sampath vuchuru · ағылшын тілі

Instant Data Scraper бойынша қазақ тіліндегі видео табылмады — сондықтан орыс және ағылшын тіліндегі видеолар ұсынылды. QR-кодты телефонмен сканерлеңіз немесе басыңыз.

Ескерту: видеолардағы мақсаттар (клиент базасын жинау т.б.) біздікі емес — біз тек құралдың батырмаларын үйренеміз және жеке деректі жинамаймыз

Негізгі ойлар

1. Турист сапарға дейін, сапар кезінде және одан кейін белсенді және енжар цифрлық іздер қалдырады
2. Деректердің үш көзі: UGC, құрылғы деректері, операция деректері
3. Жинау құралдары: API, веб-скрейпинг, мобильді деректер, іздеу сұраныстары, геотег
4. «Техникалық мүмкін» ≠ «заңды»: ҚР заңы мен сайт ережелерін тексеріңіз
5. Цифрлық із — шикізат: жалған пікір мен іріктеме ығысуын ескеріп тазалау керек



Келесі: семинар 9 — деректерді жинау және тазалау

Пайдаланылған дереккөздер

[1] Girardin F., Blat J., Calabrese F., Dal Fiore F., Ratti C. (2008). Digital footprinting: Uncovering tourists with user-generated content. IEEE Pervasive Computing, 7(4), 36–43. https://doi.org/10.1109/MPRV.2008.71

[2] Li J., Xu L., Tang L., Wang S., Li L. (2018). Big data in tourism research: A literature review. Tourism Management, 68, 301–323. https://doi.org/10.1016/j.tourman.2018.03.009

[3] Tripadvisor. 2025 Transparency Report reveals strong review submissions and improved fraud detection (пресс-релиз), 18.03.2025. https://tripadvisor.mediaroom.com/2025-03-18-Tripadvisors-2025-Transparency-Report-reveals-strong-review-submissions-and-improved-fraud-detection

[4] Instant Data Scraper: A step-by-step guide for 2026. Byteful. https://byteful.com/blog/instant-data-scraper-a-step-by-step-guide; Instant Data Scraper: Complete guide (2026). Geonode. https://geonode.com/blog/instant-data-scraper-tools

[5] Казахстан начнет считать туристов по данным сотовых операторов. Economist.kg, 06.08.2026. https://economist.kg/ekonomika/2026/08/06/kazakhstan-turisty-dannye-operatorov/

[6] Estonia: mobile positioning data for tourism statistics (кейс). GovLab Data Collaboratives. https://datacollaboratives.org/cases/estonia-mobile-positioning-data-for-tourism-statistics.html

[7] International Monetary Fund. Tourism forecasts for The Bahamas using Google Trends (2004–2018). IMF Working Paper WP/20/22, 2020. https://www.imf.org/-/media/files/publications/wp/2020/english/wpiea2020022-print-pdf.pdf

[8] «Дербес деректер және оларды қорғау туралы» ҚР 21.05.2013 № 94-V Заңы. https://adilet.zan.kz/rus/docs/Z1300000094; Morgan Lewis. Republic of Kazakhstan data protection law amendments in 2026, 14.07.2026. https://www.morganlewis.com/ru/pubs/2026/07/republic-of-kazakhstan-data-protection-law-amendments-in-2026

[9] Google Maps/Google Earth Additional Terms of Service, 2-бөлім (өзгертілген 27.01.2026). https://www.google.com/help/terms_maps

[10] Минтуризма РК переходит на цифровую статистику на базе мобильных данных. Kazinform. https://www.inform.kz/ru/minturizma-rk-perehodit-na-tsifrovuyu-statistiku-na-baze-mobilnih-dannih-6dfbf4