



D6.1 Repeatable workflow culturomics

2026-05-29

Authors: Theresa Henke, Pavel Pipek, Simon Reynaert, Sofie Meeus, María Loreto Castillo, Candida Gomes do Vale, Petr Pyšek & Ana Nova



**Funded by
the European Union**

Views and opinions expressed are those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them



D6.1 Repeatable workflow culturomics

Prepared under contract from the European Commission

Grant agreement No. 101180559

EU Horizon Europe Innovation Action

European Research Executive Agency

Project acronym: OneSTOP

Project full title: OneBiosecurity Systems and Technology for People, Places and Pathways

Project duration: 01.01.2025 – 30.06.2028 (42 months)

Project coordinator: Dr. Quentin Groom, Agentschap Plantentuin Meise (MBG)

Call: HORIZON-CL6-2024-BIODIV-01-1

Deliverable title: Repeatable workflow culturomics

Deliverable №: D6.1

WP responsible: WP6

Nature of the deliverable: Software

Dissemination level: Public

Licence of use: CC BY

Lead partner: Institute of Botany of the Czech Academy of Sciences (IBOT-CAS)

Recommended citation: Henke, T., Pipek, P., Reynaert, S., Meeus, S., Castillo, M.L., Gomes do Vale, C., Pyšek, P. & Novoa, A. (2026). *Repeatable workflow culturomics*. OneSTOP project D6.1.

Due date of deliverable: Month № 18

Actual submission date: Month № 18

Deliverable status:

Version	Status	Date	Author(s)
1.0	Draft	05 29 2026	Henke T.; EEZA-CSIC Pipek P.; IBOT-CAS Reynaert S.; MBG Meeus S.; MBG Castillo M.L.; IBOT-CAS Gomes do Vale, C.; CIBIO Pyšek P.; IBOT-CAS Novoa A.; EEZA-CSIC
1.0	Review	22 06 2026	Reyserhove, L; INBO



D6.1 Repeatable workflow culturomics



2.0	Final	25 06 2026	Kumschick, S; SU Henke T.; EEZA-CSIC Pipek P.; IBOT-CAS Reynaert S.; MBG Meeus S.; MBG Castillo M.L.; IBOT-CAS Gomes do Vale, C.; CIBIO Pyšek P.; IBOT-CAS Novoa A.; EEZA-CSIC
-----	-------	------------	--





Table of contents

Key takeaway messages.....	4
Executive summary.....	4
Non-technical summary.....	5
List of abbreviations.....	5
1. Introduction/Background.....	7
2. Objectives.....	8
3. Scope and methodological approach.....	9
4. FAIR data management and reusability.....	9
5. The workflow.....	9
5.1. Sub-workflow 1. Public interest in IAS.....	11
5.2. Sub-workflow 2. Presence of IAS in e-commerce platforms.....	18
6. Conclusions.....	26
7. Acknowledgements.....	27
8. Literature.....	27





Key takeaway messages

- Search engine query data provide a scalable proxy for analysing temporal and spatial patterns in public interest towards IAS.
- Combining the detection of unusual short-term peaks in public interest with with a contextual investigation of the digital environment during those periods can help identify the events, actors, and communication channels influencing public attention towards IAS, as well as provide insights into the sentiment underlying that attention, indicating whether interest is driven by positive or negative messages.
- Information extracted from online retail and peer-to-peer e-commerce platforms can support the monitoring of pathways facilitating the introduction and spread of IAS, by identifying species being offered for sale and informing assessments of compliance with IAS regulations.
- Standardised digital-data workflows can contribute to communication strategies, biosecurity monitoring, and future research on the human dimensions of biological invasions.

Executive summary

This deliverable presents a repeatable workflow for analysing human interest and behaviour related to invasive alien species (IAS) using digital data sources. The work was developed within Task 6.1 of the Horizon Europe OneSTOP project and focuses on two complementary dimensions of the human aspects of biological invasions: public interest in IAS, including the contextual factors shaping that attention, and the presence of IAS in e-commerce platforms.

The workflow consists of two sub-workflows. The first sub-workflow analyses online search engine query data as a proxy for public interest in IAS. Using Google search data from 2016–2025 for the 88 species included in the third update of the EU List of IAS of Union concern, the workflow identifies temporal peaks and anomalies in public attention across EU Member States and the United Kingdom. The workflow includes the definition of study parameters, data extraction, peak and anomaly detection, filtering procedures, and manual identification of potential drivers behind public interest. This approach enables the identification of events, actors, and communication channels associated with increases in public interest in IAS, while also providing context on the nature of the messages driving that attention and whether they reflect positive or negative perceptions of the species.

The second sub-workflow focuses on the presence of IAS in online trade and peer-to-peer e-commerce platforms. Two surveys conducted in 2021 and 2024 assessed the availability of IAS of Union concern on e-commerce platforms in six EU Member States and internationally operating platforms. The workflow includes the definition of study parameters, identifying suitable e-commerce platforms, systematically searching for IAS using scientific and common names, recording listing characteristics, and analysing temporal trends in species presence. Results indicate that many IAS of Union Concern remain available to purchase online despite regulatory restrictions, highlighting the importance of monitoring digital trade pathways, and border biosecurity.





D6.1 Repeatable workflow culturomics

Together, these workflows provide transparent, standardised, and reusable methodologies for studying the human dimensions of biological invasions using digital data. The approaches developed here can support future research, improve communication and outreach strategies, and contribute to policy implementation and biosecurity monitoring within and beyond the OneSTOP (OneBiosecurity Systems and Technology for People, Places and Pathways) project.

Non-technical summary

Invasive alien species (IAS) are plants and animals introduced outside their natural range that can harm biodiversity, ecosystems, economies, and human well-being. Human activities play a central role in the introduction and spread of these species, but people are also essential for preventing and managing IAS. Understanding how the public interacts with and responds to IAS is therefore an important part of improving biosecurity and conservation efforts.

This deliverable describes a repeatable workflow developed within the OneSTOP project to study public interest and online trade related to IAS using digital information sources. The workflow uses publicly available online data to better understand how people search for, discuss, perceive, and trade IAS across Europe.

The first part of the workflow analyses online search activity using Google search data. By examining changes in search interest over time, the workflow can identify moments when public attention towards certain IAS increases significantly. These peaks in attention are then linked to possible causes such as media reports, management actions, new invasions, or health and environmental concerns. This contextual information helps reveal whether increased interest is associated with positive or negative messages and perceptions surrounding the species.

The second part of the workflow investigates the presence of IAS in online trade. Surveys carried out in 2021 and 2024 examined whether species listed, or proposed for listing, as IAS of Union concern were available on online retail and peer-to-peer trading platforms. The results showed that many IAS were still being sold online, despite existing EU regulations prohibiting their trade. For most of the IAS recorded as present in e-commerce information on where it is traded from were unavailable. However, where information on the origin was retrieved, it indicated that across both years, IAS were sold and imported from inside and outside of the EU.

The workflows developed in this deliverable provide a standardised and reusable method for collecting and analysing digital data related to IAS. These approaches can help researchers, policymakers, and practitioners better understand public behaviour, improve communication campaigns, monitor compliance with regulations, and support future biosecurity actions across Europe.

List of abbreviations

IAS - Invasive alien species
API - Application Programming Interface
EU - European Union
MS - (EU) member state





Country abbreviations:

AT - Austria
BE - Belgium
BG - Bulgaria
CY - Cyprus
CZ - Czech Republic
DE - Germany
DK - Denmark
EE - Estonia
ES - Spain
FI - Finland
FR - France
GB - Great Britain
GR - Greece
HR - Croatia
HU - Hungary
IE - Ireland
IT - Italy
LT - Lithuania
LU - Luxembourg
LV - Latvia
MT - Malta
NL - Netherlands
PL - Poland
PT - Portugal
RO - Romania
SE - Sweden
SI - Slovenia
SK - Slovakia





1. Introduction/Background

Biological invasions are a global phenomenon of increasing severity and are widely recognized as one of the leading threats to biodiversity (IPBES 2023). Invasive alien species (IAS) can cause significant environmental (Kumschick et al. 2020) and socio-economic (Bacher et al. 2018) impacts, affecting ecosystem functioning, native biodiversity and human well-being (Vilà et al. 2011, Pyšek et al. 2012, Gallardo et al. 2016, Shackleton et al. 2019a, IPBES 2023).

Biological invasions are inherently shaped by human activities. Humans act as key drivers of species introductions (both intentionally and unintentionally), and are directly affected by the consequences of invasions (Shackleton et al. 2019b). Furthermore, societal responses, including management actions and policy implementation, play a critical role in determining invasion outcomes (Novoa et al. 2025). As a result, there is growing recognition of the importance of understanding the human and social dimensions of biological invasions, including public interest and behaviours (Shackleton et al. 2019b).

Public interest in IAS is particularly important in this context. Individuals who are aware of IAS-related risks and understand their impacts are more likely to support and comply with management measures (Novoa et al. 2017a). As such, effective communication and outreach have been an essential part of IAS research and management strategies (Pocock et al. 2024). But designing effective and targeted outreach and communication campaigns requires an understanding of the dynamics of the public's interest and attention, as well as the nature of the messages and perceptions underlying that attention (Davis et al. 2018).

The recently established scientific field of conservation culturomics (Ladle et al. 2016, Jarić et al. 2021) provides opportunities to address this challenge. Conservation culturomics leverages large-scale publicly available digital data (such as relative search volumes, social media content and online trade information) to analyse relationships between human and nature (Novoa et al. 2025). While these tools come with their own biases that require addressing, they have increasingly been applied in conservation science to investigate public interest in biodiversity, track emerging environmental issues, and better understand human–nature interactions at broad spatial and temporal scales (Ladle et al. 2016).

Given the importance of incorporating human and social dimensions when addressing biological invasions, generating timely data at larger scales becomes crucial. Developing repeatable workflows can increase the speed and scale of data collection by providing transparent methodologies, comparable and standardized outcomes, and greater opportunities for collaboration (Reiter et al. 2021, Groom et al. 2025).

Within the Horizon Europe project **OneSTOP** (OneBiosecurity Systems and Technology for People, Places and Pathways), Task 6.1 focuses on applying such approaches to decipher public interest (e.g. online attention) and behaviour (i.e. trade activities) towards IAS using





D6.1 Repeatable workflow culturomics

digital data sources. In addition to identifying temporal patterns in public attention, these approaches can provide contextual information on the events, actors, communication channels, and positive or negative messages associated with changes in public interest. The development of a workflow in this deliverable was conducted on the basis of two case studies focusing on species included in the List of IAS of Union Concern (henceforth Union List). This list, established under the EU IAS regulation (1143/2014) on the prevention and management of the introduction and spread of invasive alien species, represents a core policy instrument of the EU for addressing biological invasions. It comprises species that are alien to the EU, pose significant environmental risks, and have undergone formal risk assessment procedures. Following the regulation's adoption in October 2014, the initial list was published in 2016 and has been regularly updated, expanding from 37 species (23 animals and 14 plants) to 88 species (47 animals and 41 plants, including both aquatic and terrestrial species) following the 3rd update in 2022 (European Commission: Directorate-General for Environment 2023). As the OneSTOP project commenced in early 2025 (i.e. before the fourth update of the Union List), the analyses presented in this deliverable are based on the 88 species included in the 3rd update of the Union List.

The two case studies underpinning this deliverable address complementary aspects of the human dimensions of biological invasions: (i) public interest in IAS as reflected in online search activity and the contextual factors shaping that attention and (ii) the presence and dynamics of IAS in e-commerce as a proxy for potential introduction pathways. Together, these case studies form the basis for the development and evaluation of a repeatable workflow designed to retrieve, process, and analyse digital data related to the human dimensions of IAS. This workflow contributes directly to the objectives of Task 6.1 (see below) and provides a methodological foundation for subsequent activities within WP6 (i.e., supporting EU Regulation 1143/2014 and providing advice for Target 6 of the Global Biodiversity Framework), as well as communication and engagement strategies.

2. Objectives

The overall objective of this deliverable is to develop and document a repeatable workflow for analysing human interest and behaviours towards IAS using digital data sources. More specifically, this deliverable aims to:

- a) Develop a workflow to quantify public interest in IAS and investigate the events, actors, communication channels, and positive or negative messages associated with changes in public attention.
- b) Develop a workflow to assess IAS presence in e-commerce and their role as pathways for introduction and spread.
- c) Provide methodologies that can be reused and further developed beyond OneSTOP

Due to the different nature and intricate details of aims a and b, we divided the workflow into two separate sub-workflows.





3. Scope and methodological approach

Task 6.1 was originally conceived to explore how social media and other digital data sources could be used to understand public responses to IAS, including public attitudes and perceptions. Since the preparation of the project proposal, the social media landscape has changed substantially, affecting data accessibility, reproducibility, and comparability across platforms and countries (see Heiss and Freiling 2026, and Deliverable 2.3). As a result, and following discussions within the Task 6.1 team and with the broader consortium during the OneSTOP annual meeting in Almería in January 2026, it was decided to focus on two complementary and scalable digital data sources: search engine query data and e-commerce information. Search-query data provide a proxy for public attention and information-seeking behaviour, while e-commerce data provide insights into behaviours associated with IAS introductions and compliance with IAS legislation. These approaches enable the identification of events, topics, actors, and communication channels associated with changes in public attention towards IAS, while also providing contextual information on the drivers underlying that attention and whether they reflect positive or negative perceptions of the species. In addition, they provide information on behaviours relevant to IAS introduction pathways. Together, they contribute to the original objectives of Task 6.1 by improving understanding of the human dimensions of biological invasions and by supporting subsequent activities within WP6, including Task 6.2.

4. FAIR data management and reusability

The workflows developed in this deliverable were designed according to Horizon Europe Open Science principles and aim to maximise transparency, reproducibility, and reuse. Workflow documentation, analytical code, data templates, and supporting materials are available through the [OneSTOP GitHub repository](#) and archived through Zenodo (<https://zenodo.org/records/20560588> and <https://zenodo.org/records/20561150>). Some underlying data (i.e. names of the e-commerce platforms, identity of people involved in peer-to-peer trading platforms, etc.) are subject to restrictions associated with ethical considerations. By combining openly available workflows, analytical code and metadata records, the approaches presented in this deliverable are intended to be reusable both within and beyond the OneSTOP project.

5. The workflow

This deliverable presents two complementary sub-workflows designed to analyse different dimensions of human–IAS interactions using digital data sources. This section describes these two sub-workflows (Fig. 1).



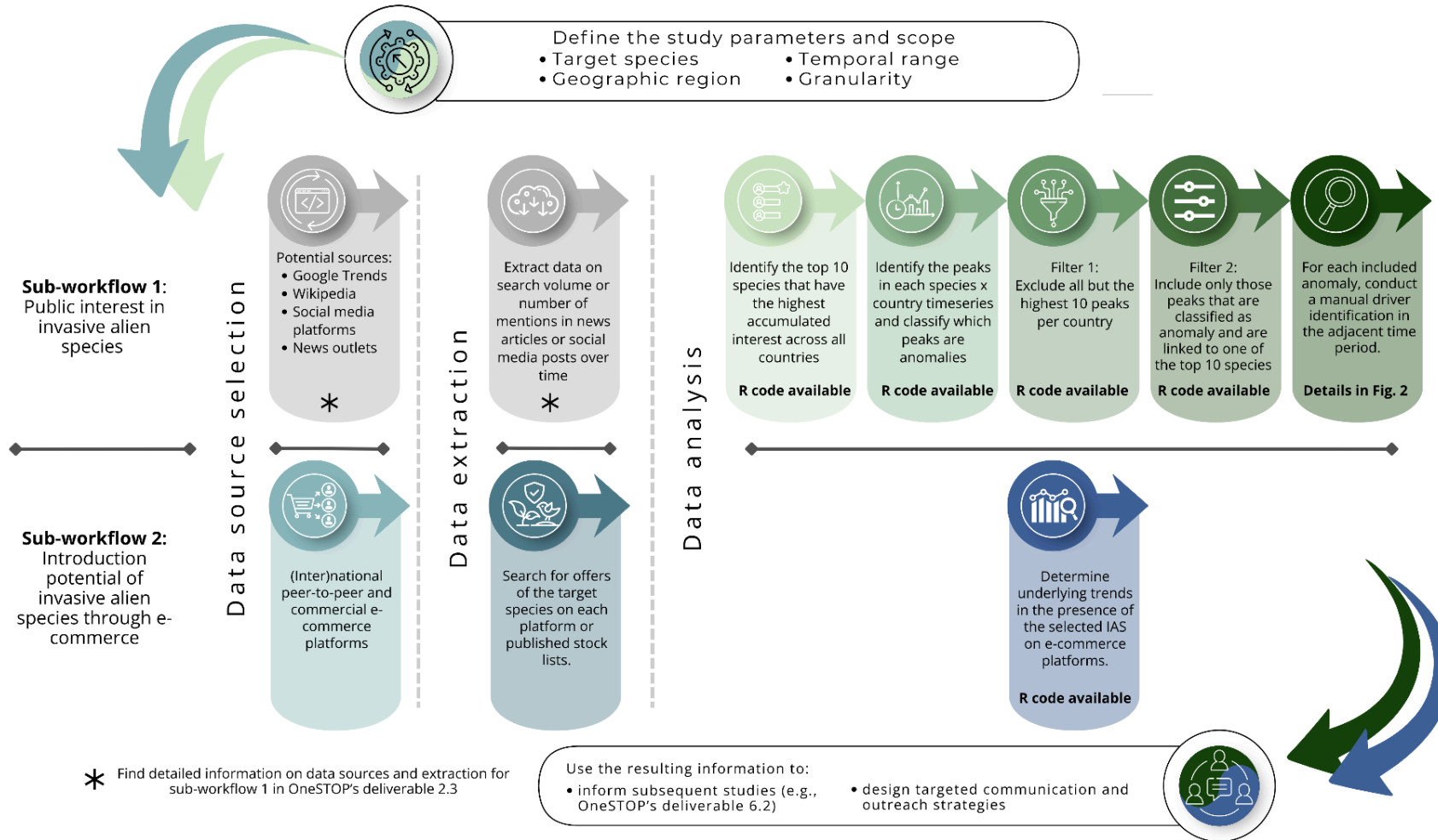


Figure 1. Schematic overview of the repeatable workflow. The workflow is divided into sub-workflow 1 addressing public interest in IAS and the contextual factors underlying changes in public attention, and sub-workflow 2 addressing potential introduction of IAS through e-commerce. Both sub-workflows have a shared first step (i.e. defining the study parameters and scope) and last step (i.e. Using the resulting information to inform subsequent studies and communication strategies). Note that step 2 and 3 of sub-workflow 1 are detailed in OneSTOP deliverable 2.3





5.1. Sub-workflow 1. Public interest in IAS

This sub-workflow analyses search engine query data as a proxy for public interest in IAS. Search engine query data provides insights into what people are actively seeking information about, allowing researchers to track temporal and spatial dynamics of public interest on particular topics. In recent years, the use of search engine query data as a proxy for interest has emerged as a powerful tool in the context of Conservation Culturomics (Scharnow and Vogelgesang 2011, Nghiem et al. 2016). These approaches have been successfully applied in conservation science to analyse attention towards diverse biodiversity and environmental issues (Correia et al. 2019).

Here, we developed a sub-workflow to systematically gather and analyse search engine query data related to IAS and identify underlying drivers and contextual factors shaping public interest. Overall, sub-workflow 1 consists of nine main steps: (1) Define parameters and scope of the study, (2) select the suitable source of data, (3) extract the data (see deliverable 2.3 for more details), (4) identify the top 10 species (this number can be adjusted to the specific aims of the project) with the highest accumulated public interest, (5) detect peaks and classify anomalies, (6) filter data to include the top 10 peaks per country (this number can be adjusted to the specific aims of the project), (7) filter data for those peaks that are classified as an anomaly and are linked to one of the target species, (8) conduct a manual driver identification, and (9) utilize the results to inform subsequent projects and communication strategies (Fig. 1).

The first step involves defining the scope of the analysis (step 1). This includes selecting the target species, defining the geographical region of interest, and determining the temporal range and resolution of the data. Once the parameters are defined, a suitable data source is selected (step 2). Potential sources could be Google Trends, social media platforms, Wikipedia, or news outlets etc. In particular, Google Trends provides aggregated search interest data that can be structured into time-series representing public attention over time (Alam and Hulme, 2026).

The next steps (Steps 3 and 4) include data extraction (see OneSTOP's Deliverable 2.3 for details on this step) and the identification of the species with the highest accumulated interest values, respectively. The extracted data are subsequently processed to construct species–country time-series, which form the basis for the peak and anomaly analysis.

The peak and anomaly analysis (Step 5) aims to identify peaks in public interest (i.e., peaks in activity), and detect anomalies by removing underlying seasonal and long-term trends, allowing for the identification of unusual increases in attention that may be linked to specific events. To reduce the number of cases for further analysis, peaks can be filtered based on their prevalence, whether they constitute anomalies, and the species they are linked to (steps 6 and 7).

To better understand the drivers behind the observed patterns, a manual identification (step 8) is included (see Fig. 2). In this step, for each peak classified as an anomaly and linked to one of the 10 species with the highest accumulated public interest, contextual information is collected





D6.1 Repeatable workflow culturomics

through targeted Google searches, focusing on the time period surrounding the peak. The content is then classified based on the type of information, the platform on which it was shared, the actor groups involved, and the nature of the messages associated with the increased attention, including whether they reflect positive or negative perceptions of the species.

The sub-workflow ultimately produces a time-series of public interest, identifies significant peaks and anomalies, and provides insights into the potential drivers and framing of public attention. In a final step, these outputs can support the development of targeted communication strategies and improve understanding of how public awareness and perceptions of IAS evolve over time (see OneSTOP's deliverable 6.2. for an example).

To demonstrate this approach, the following case study identifies and analyzes the drivers of public interest in invasive alien species included in the Union List by detecting search query anomalies across Europe and by manually investigating the specific events, topics, and actors responsible for those digital spikes:

- **Step 1:** The parameters and scope of the case study were established by defining the target species, geographic coverage, temporal range and granularity (i.e. whether the data will be extracted on a daily, weekly, or monthly resolution). The target species list encompassed all 88 species included in the third update of the Union List published in 2022 (European Commission: Directorate-General for Environment 2023). The temporal range was set from January 2016 to December 2024 with data being collected at a weekly resolution. The geographic coverage of the case study comprised all current EU member states (MS) as well as the United Kingdom as former EU MS as the timeline covers segments prior to the UK's exit.
- **Step 2 and 3:** We selected Google as the source of our search query data. We obtained the proportional weekly Google search volume for each of the target species in each of the target countries from January 1, 2016, until December 31, 2024, relative to all Google searches in each country during that period. To do this, first, we used Google topics to include results across searches made with synonyms, typos, and similar or related terms across EU languages related to each species name (Cooper et al. [2019](#)). Then, we used Google Health Trends API (application program interface) and a dedicated Python library (Python version 3.8; google-API-python-client; <https://pypi.org/project/google-api-python-client/>) to extract Google search volumes for each species/topic (see Vardi et al. [2021](#) for details on data extraction). It is worth noting that data extracted from Google Trends could provide comparable information to that extracted from Google Health Trends (see Alam and Hulme, 2026). However, compared with the relative search volumes provided by Google Trends, the Google Health Trends API returns a probability-based metric that is not rescaled within each query period, which allows more consistent comparison of search data extracted across different time intervals and thus offers methodological advantages for longitudinal analyses (Neumann et al. 2023).





- Step 4:** Aiming to filter data in following steps (see below), we identified the top 10 species that received the highest accumulated public interest between 2016 and 2022 across all 28 target countries. Specifically, the top 10 species included *Procyon lotor*, *Myocastor coypus*, *Vespa velutina nigrithorax*, *Nyctereutes procyonides*, *Ondathra zibethicus*, *Heracleum mantegazzianum*, *Heracleum sosnowskyi*, *Impatiens glandulifera*, *Ailanthus altissima*, and *Nasua nasua* (R code in this [Github repository](#)).
- Step 5:** We identified peaks across all species x country trajectories and subsequently detected which of those peaks were classified as anomalies. To ensure reproducibility, peak identification and anomaly detection were performed using a standardised analytical procedure implemented in R (R code in this [Github repository](#)). Peaks were first identified automatically within each species-country time series using the `findpeaks()` function of the *pracma* R package (Borchers 2025). To distinguish unusual increases in public interest from recurring seasonal patterns and long-term trends, anomaly detection was subsequently applied to the time series. Utilizing the *anomalize* R package (Dancho and Vaughan 2023), we first decomposed each time series into its seasonal, trend, and remainder components using the `time_decompose()` function, then performed an anomaly detection on the remainder component using `anomalize()`, and subsequently recomposed the timeseries with `time_recompose()`. This procedure allowed the identification of observations that deviated substantially from expected baseline levels of public interest.
- Step 6:** We identified the 10 highest peaks per country (i.e., those presenting the highest values of relative search interest). This resulted in a total of 280 peaks (see the [Github repository](#) for details).
- Step 7:** Overall, 153 out of the 280 peaks identified in step 6 were classified as anomalies and linked to the species with the highest public interest (step 4). These were included for further analysis (Table 1).

Table 1: Overview of species included in further analysis and the respective number of peaks found among the top 10 peaks per country, the number of anomalies included in further analysis, and the countries these peaks and anomalies were detected in. Please note that there was no difference between countries where peaks and anomalies occurred per species.

Species	Common name	n peaks among top 10 peaks per country	n anomalies included in further analysis	Countries
<i>Procyon lotor</i>	Raccoon	127	71	26 (AT, BE, BG, CY, CZ, DE, DK, EE, FI, FR, GB, GR, HR, HU, IE, IT, LT, LU, LV, MT, NL PL RO, SE, SI, SK)





<i>Myocastor coypus</i>	Coypu/Nutria	26	21	11 (CZ, EE, FI, HR, IE, IT, LT, PL, RO, SI, SK)
<i>Vespa velutina nigrithorax</i>	Asian hornet	52	25	12 (AT, BE, CZ, ES, FR, GB, HR, IE, NL, PT, SE, SK)
<i>Nyctereutes procyonides</i>	Raccoon dog	20	15	10 (CZ, DK, EE, FI, GB, HR, LT, LU, PL, SE)
<i>Ondatra zibethicus</i>	Muskrat	0	0	0
<i>Heracleum mantegazzianum</i>	Giant hogweed	9	9	6 (HR, HU, IE, SE, SI, SK)
<i>Heracleum sosnowskyi</i>	Sosnowskyi's hogweed	6	4	2 (LT, PL)
<i>Impatiens glandulifera</i>	Himalayan balsam	5	5	4 (EE, LV, SE, SI)
<i>Ailanthus altissima</i>	Tree of heaven	2	2	2 (AT, HR)
<i>Nasua nasua</i>	South American coati	1	1	1 (SE)

- Step 8:** For each of these 153 peaks, we conducted a manual search on Google search engine for the adjacent time period (one week preceding and one week following the peak) using scientific names, synonyms, and species' common names in the respective local language. Using a standard template (see the [Github repository](#)), we recorded information on the context in which the target species was mentioned, including the topics covered, the platforms involved, the actor groups responsible for the content, and the overall framing of the information. We identified whether the content covered one or multiple specific topics (Fig. 2). Subsequently, we classified the identified specific topics more detailed into broad topics such as impacts on health/recorded deaths, first reports, and approaching arrival. We then further documented the type of platform the content was shared on and classified the associated actor group. For the actor groups we considered news journalists (regional/national/international), management professionals (regional/national/international), other experts, management services, non-commercial interest groups, and lay people. For consistency, information was recorded using pre-designed categories for the manual search (Fig. 2). Media reporting peaks were unevenly distributed across species and countries, with some IAS generating peaks in multiple countries while others were associated with relatively few (Fig. 3). Confidence scores were generally moderate to high where peaks were identified, indicating consistent evidence for the detected events. Low confidence scores were particularly abundant in peaks linked to *Procyon lotor* and *Myocastor coypus* (Fig. 3). Similarly, the topics associated with increased public interest differed among species, although





management-related topics were common across all taxa. (Fig. 4). Increased public interest in the four plant species was mostly linked to management, (scientific and conservation) projects targeting the species, and reported health concerns, with additional reporting of first observations in countries and/or new regions linked specifically to *Heracleum mantegazzianum* (Fig. 4). Meanwhile, the range of topics showed much higher variation for *Vespa velutina nigrithorax* and the mammal species included in the analyses, with the exception of *Nasua nasua*, where only one peak was analysed (Fig. 4).

A summary of the driver-identification results is presented here because it provides important context for interpreting anomalies in public interest and demonstrates how the workflow can elucidate the events, actors, and messages associated with changes in public attention, including whether such attention is driven by positive or negative narratives surrounding IAS. Detailed analyses will be presented elsewhere (Henke et al., *in preparation*).

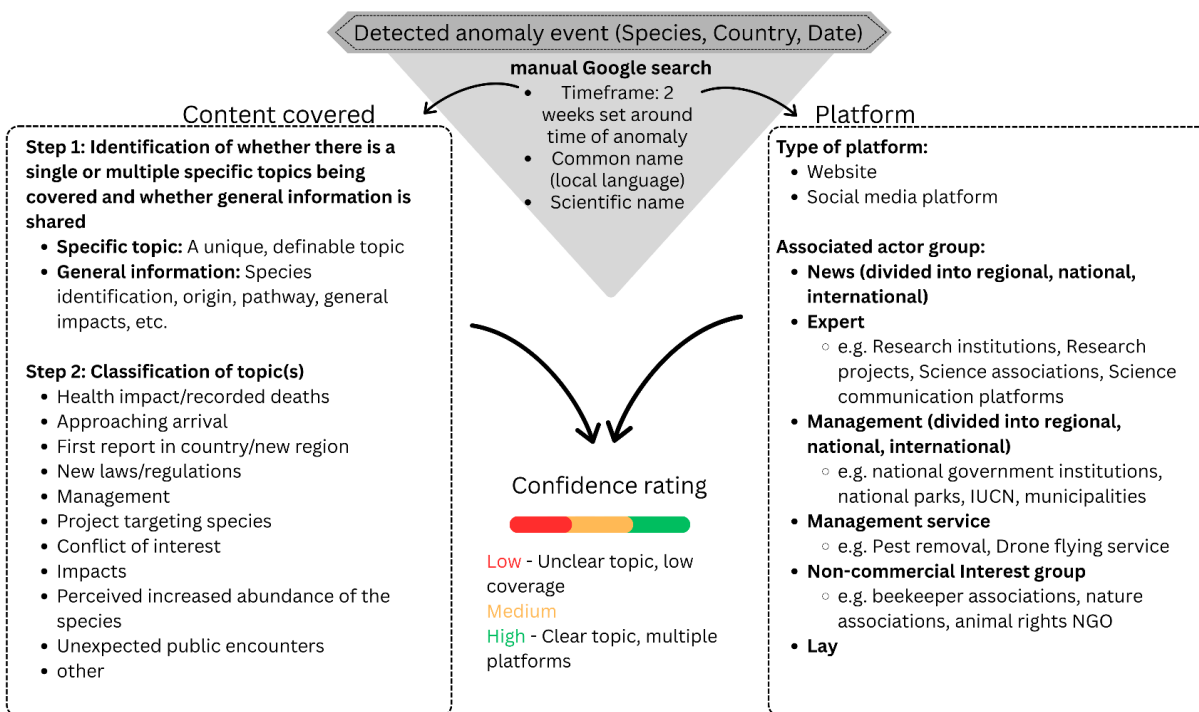


Figure 2. Schematic overview for the manual identification of potential drivers of peaks identified in public interest in species listed on the Union list.



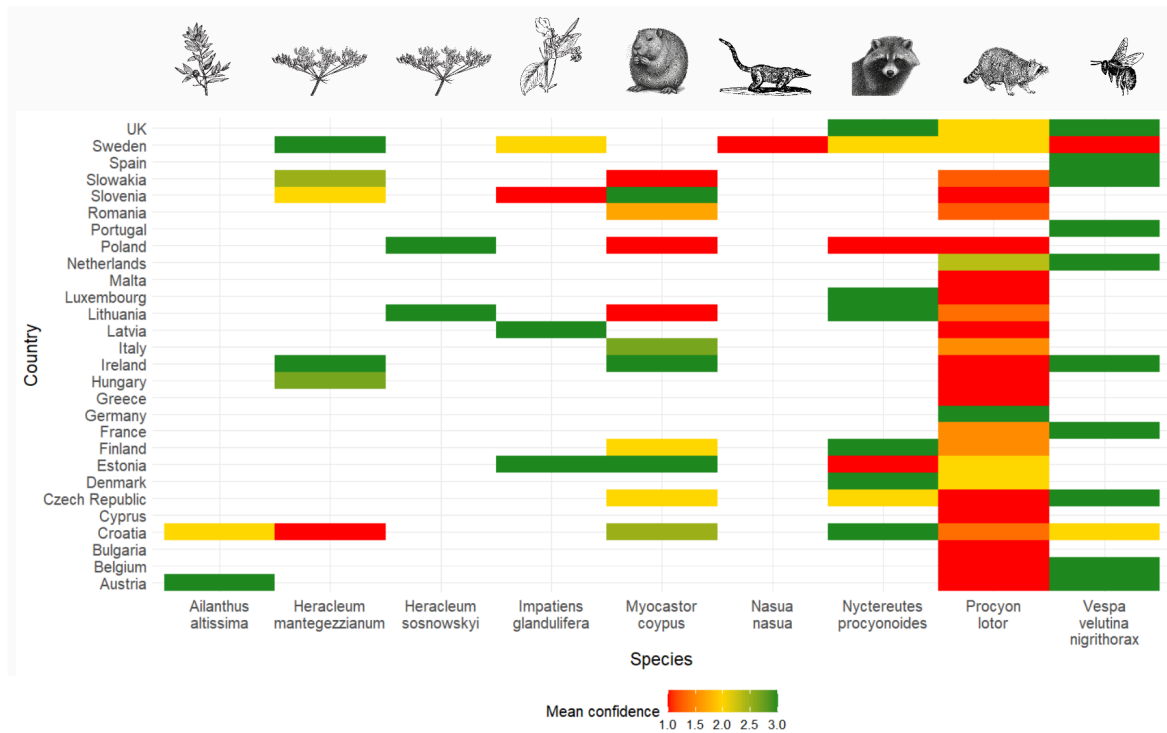


Figure 3. Mean confidence scores for species-specific media reporting peaks across MS. Confidence scores reflect the level of certainty associated with the identification of drivers of increased public interest in selected IAS. Note that while *Ondathra zibethicus* was among the ten species with the highest accumulated public interest, it was not linked to any of the analysed peaks. Darker green shading indicates higher mean confidence, while grey cells indicate the absence of detected peaks.



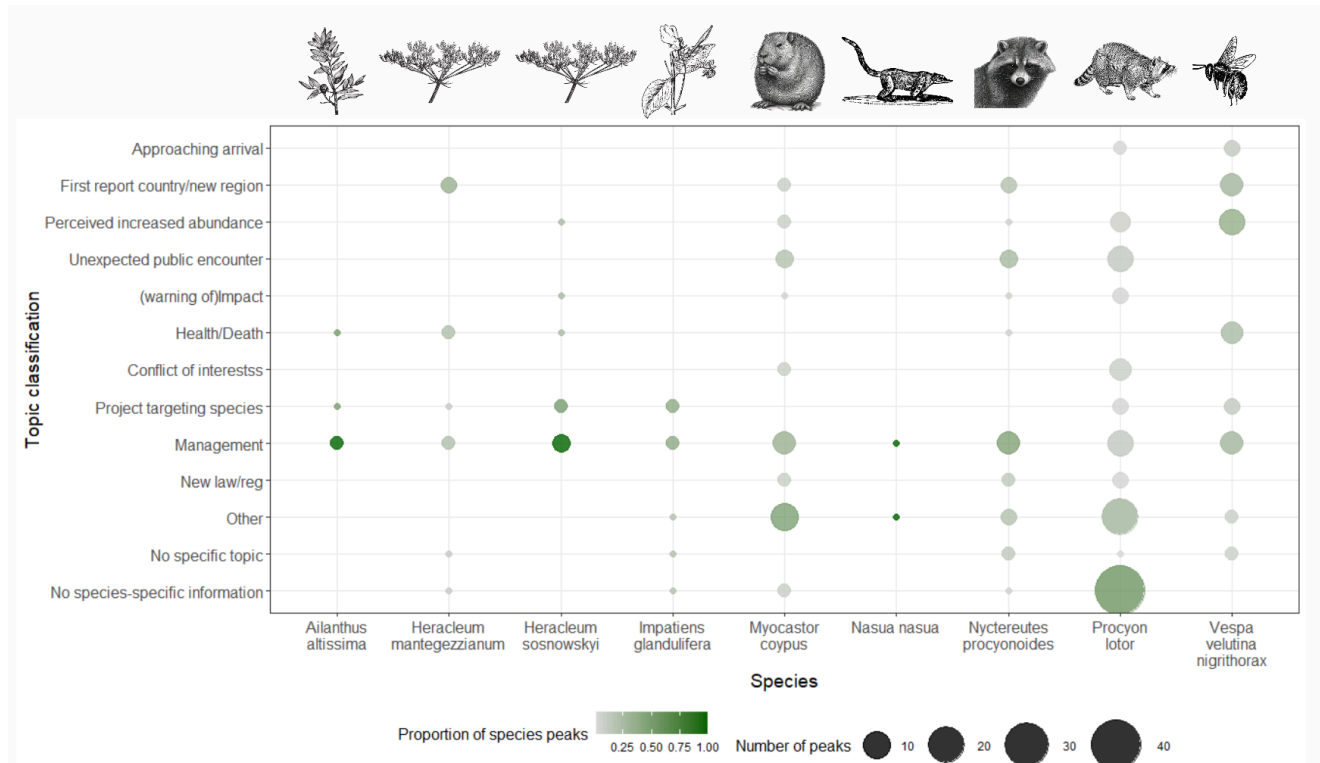


Figure 4. Distribution of topic classifications associated with increased public interest in selected IAS. Note that while *Ondathra zibethicus* was among the ten species with the highest accumulated public interest, it was not linked to any of the analysed peaks. Bubble size represents the number of peaks assigned to each topic-species combination, while colour intensity indicates the proportion of all peaks for a given species represented by that topic. Larger, darker circles therefore denote topics that were both frequent and relatively important within a species' media coverage.

- Step 9:** The outputs of this workflow informed OneSTOP Task 6.2 by providing information on the relative levels of public interest associated with different IAS. These results contributed to the selection of species included in subsequent public perception surveys and provided contextual information for interpreting differences in public awareness and engagement across taxa (see Deliverable 6.2.).In addition, the identification of the events, actors, and messages associated with peaks in public attention helped provide insight into the positive and negative narratives surrounding different species, thereby supporting the design and interpretation of subsequent activities within Task 6.2.





5.2. Sub-workflow 2. Presence of IAS in e-commerce platforms

Under the EU Regulation 1143/2014, Member States are required to prohibit the sale and distribution of species listed on the Union list (European Commission: Directorate-General for Environment 2023). However, the effectiveness of these measures depends on their implementation and enforcement, which may vary across countries and platforms.

Ornamental and pet trade represent some of the most important pathways for the introduction and spread of IAS (van Kleunen et al. 2018). Monitoring the presence of IAS in e-commerce thus provides an opportunity to assess compliance and evaluate the effectiveness of regulatory measures.

Sub-workflow 2, developed to assess IAS presence in e-commerce platforms, follows a structured and repeatable process, consisting of five main steps: (1) Define parameters and scope of the study, (2) Identify peer-to-peer and commercial e-commerce platforms, (3) Extract data on offers in each platform, (4) Determine trends in the presence of IAS in e-commerce platforms, and (5) utilize the results to inform subsequent projects, legislation and communication strategies (Fig. 1).

The first step, shared with sub-workflow 1 described above, involves defining the scope of the analysis, including the selection of species, and geographic regions. This ensures that the study is aligned with the research objectives and captures relevant variation across markets. In step 2, suitable e-commerce platforms should be identified using standardised search terms and expert knowledge (e.g., Novoa et al. 2017b). These include both online retail (including physical stores that also sell online), and peer-to-peer trading platforms.

Once the platforms are identified, each is systematically searched for the presence of target species (step 3). Searches are conducted using scientific names, synonyms, and common names in the relevant languages. Where available, platform-specific search tools are used; otherwise, manual inspection of stock lists is performed. Data is then collected manually (see steps below) using a standard template (see the [Github repository](#)) to ensure consistency across surveyors and platforms. Information recorded should include species name, availability, characteristics of the listing, and pricing, when applicable. For each observation, when possible, the post should be checked for consistency between the provided species name and the species displayed in the sales picture.

In step 4, the collected data are analysed to identify temporal trends in species presence (find R code in this [Github repository](#)). Species are categorised based on their occurrence across survey years, for example as “entry” (i.e., not targeted or found in e-commerce in 2021, but present in 2024), “persistent” (i.e., found in e-commerce in 2021 and 2024), or “exit” (i.e., found in e-commerce only in 2021). These categories provide insights into changes in trade dynamics and potential effects of regulatory measures. The results of these analyses are beyond the scope of this deliverable and will be published in Henke et al. (in preparation).





D6.1 Repeatable workflow culturomics

In a final step, the outputs of this analysis can support the development of targeted communication strategies and improve our understanding of compliance with IAS regulations.

We showcase the steps of sub-workflow 2 on the basis of two e-commerce surveys, conducted in 2021 and 2024. These surveys were originally implemented as part of previous work undertaken for the European Commission in collaboration with IUCN to support the evaluation of the implementation of Regulation (EU) No 1143/2014. The data had not previously been published and permission was granted for their use within OneSTOP. Within the OneSTOP project, these datasets served as case studies for the development, documentation, standardisation, and evaluation of a repeatable workflow for monitoring IAS presence in e-commerce environments. Consequently, the principal output of Task 6.1 is the workflow itself rather than the generation of new survey data.

Specifically, these surveys investigated the introduction potential of IAS via online trade, specifically focussing on the IAS included or proposed for inclusion in the Union List. The surveys were conducted on a set of online trade and peer-to-peer platforms based in six EU member states: Czech Republic, France, Germany, Italy, the Netherlands, and Spain as well as a set of internationally acting ones. Member states were selected based on a previously undertaken trade data analysis of potential IAS into and within the EU, that evaluated trade volume, value, trends, and trade flow (IUCN 2018). The selected MS represented those countries with the highest trade numbers for mammals, ornamental fish, reptiles and horticultural plants. Spain was included to account for the four countries with the highest population numbers next to France, Germany, and Italy.

- **Step 1:** The parameters and scope of the case study were set by determining the temporal range of the surveys as well as a list of species and countries to be included. The surveys were set in 2021 and 2024, each a year prior to the publishing of the third and fourth update of the Union List, respectively (European Commission: Directorate-General for Environment 2023). The list of target species was designed to include IAS already listed on the Union List at the time and those proposed for inclusion in the upcoming update of the List. To efficiently concentrate effort, species were excluded from the surveys if their known invasions are solely linked to unintentional introduction pathways and if they are unlikely to be found on e-commerce platforms. In the end, the list of target species included 81 IAS (43 animal and 38 plant species) for the 2021 survey and 100 IAS (53 animal and 47 plant species) for the 2024 survey (Table 2). In both years, the surveys aimed to sample 50 online retailers (representing 25 horticulture and 25 pet/aquarium trading platforms) as well as 10 peer-to-peer trading platforms per member state. Additionally, the study aimed to sample 10 international horticultural and 10 international pet/aquarium platforms.





Table 2. List of plant and animal species included in the 2021 and 2024 surveys. The variables presented are binary indicators. The variables surveyed_2021 and surveyed_2024 indicate whether a species was included in the list of species targeted by the respective survey during the respective year (1 = included; 0 = not included). The variables e_commerce_presence_2021 and e_commerce_presence_2024 indicate whether a species was detected at least once among the surveyed e-commerce platform during the respective year (1 = detected; 0 = not detected).

Species	surveyed_2021	e_commerce_presence_2021	surveyed_2024	e_commerce_presence_2024
Plant species				
<i>Acacia mearnsii</i>	0	0	1	1
<i>Acacia saligna</i>	1	1	1	1
<i>Ailanthus altissima</i>	1	1	1	1
<i>Alternanthera philoxeroides</i>	1	1	1	0
<i>Andropogon virginicus</i>	1	1	1	0
<i>Asclepias syriaca</i>	1	1	1	1
<i>Baccharis halimifolia</i>	1	1	1	1
<i>Broussonetia papyrifera</i>	0	0	1	0
<i>Cabomba caroliniana</i>	1	1	1	1
<i>Cardiospermum grandiflorum</i>	1	1	1	1
<i>Celastrus orbiculatus</i>	1	1	1	1
<i>Cortaderia jubata</i>	1	1	1	0
<i>Cortaderia selloana</i>	0	0	1	0
<i>Crassula helmsii</i>	0	0	1	0
<i>Delairea odorata</i> (syn. <i>Senecio mikanioides</i>)	0	0	1	1
<i>Ehrharta calycina</i>	1	0	1	0
<i>Eichhornia</i> (<i>Pontederia</i>) <i>crassipes</i>	1	1	1	1
<i>Elodea nuttallii</i>	1	1	1	1
<i>Gunnera tinctoria</i>	1	1	1	1
<i>Gymnocoronis spilanthoides</i>	1	1	1	1
<i>Hakea sericea</i>	1	0	1	0
<i>Heracleum mantegazzianum</i>	1	1	1	0
<i>Heracleum persicum</i>	1	0	1	0
<i>Heracleum sosnowskyi</i>	1	0	1	0
<i>Humulus scandens</i> (syn. <i>Humulopsis scandens</i>)	1	1	1	1
<i>Hydrocotyle ranunculoides</i>	1	1	1	1
<i>Impatiens glandulifera</i>	1	1	1	1
<i>Koenigia polystachya</i> (syn. <i>Persicaria wallichii</i> ; <i>Polygonum polystachyum</i>)	1	1	1	1
<i>Lagarosiphon major</i>	1	1	1	1
<i>Lespedeza cuneata</i> (syn. <i>Lespedeza cuneata</i> var. <i>cuneata</i>)	1	1	1	0
<i>Ludwigia grandiflora</i>	1	0	1	0
<i>Ludwigia peploides</i>	1	1	1	1
<i>Lygodium japonicum</i>	1	1	1	0
<i>Lysichiton americanus</i>	1	1	1	1
<i>Myriophyllum aquaticum</i>	1	1	1	1



D6.1 Repeatable workflow culturomics



<i>Myriophyllum heterophyllum</i>	1	1	1	1
<i>Pennisetum setaceum</i> (syn. <i>Cenchrus setaceus</i>)	1	1	1	1
<i>Phytolacca americana</i>	1	1	0	0
<i>Pistia stratiotes</i>	1	1	1	1
<i>Prosopis juliflora</i> (syn. <i>Neltuma juliflora</i>)	1	1	1	1
<i>Pueraria lobata</i> (montana)	1	1	1	1
<i>Reynoutria japonica</i> (syn. <i>Fallopia japonica</i>)	0	0	1	1
<i>Reynoutria sachalinensis</i> (syn. <i>Fallopia sachalinensis</i>)	0	0	1	1
<i>Reynoutria x bohémica</i> (syn. <i>Fallopia x bohémica</i>)	0	0	1	1
<i>Salvinia molesta</i>	1	1	1	0
<i>Tradescantia fluminensis</i>	0	0	1	1
<i>Triadica sebifera</i>	1	0	1	1
<i>Zostera japonica</i>	0	0	1	0
Animal species				
<i>Acridothères cristatellus</i>	0	0	1	1
<i>Acridothères tristis</i>	1	1	1	0
<i>Alopochen aegyptiacus</i>	1	1	1	1
<i>Ameiurus melas</i>	1	1	1	1
<i>Ameiurus nebulosus</i>	1	1	0	0
<i>Axis axis</i>	1	1	1	0
<i>Brachyponera chinensis</i>	0	0	1	0
<i>Callosciurus erythraeus</i>	1	0	1	0
<i>Callosciurus finlaysonii</i>	1	1	1	1
<i>Channa argus</i>	1	1	1	1
<i>Cherax destructor</i>	0	0	1	1
<i>Cherax quadricarinatus</i>	0	0	1	1
<i>Cipangopaludina chinensis</i> (syn. <i>Bellamya chinensis</i>)	0	0	1	0
<i>Faxonius immunis</i>	0	0	1	0
<i>Fundulus heteroclitus</i>	1	1	1	0
<i>Gambusia affinis</i>	1	1	1	1
<i>Gambusia holbrooki</i>	1	1	1	1
<i>Herpestes javanicus</i> (syn. <i>Herpestes auropunctatus</i>)	1	0	1	0
<i>Lampropeltis getula</i>	1	1	1	1
<i>Lepomis gibbosus</i>	1	1	1	1
<i>Lithobates catesbeianus</i>	1	0	1	0
<i>Marisa cornuarietis</i>	0	0	1	1
<i>Misgurnus anguillicaudatus</i>	0	0	1	1
<i>Misgurnus bipartitus</i>	0	0	1	0
<i>Muntiacus reevesi</i>	1	0	1	0
<i>Myiopsitta monachus</i>	0	0	1	1
<i>Myocastor coypus</i>	1	1	1	1
<i>Nasua nasua</i>	1	1	1	0
<i>Neogale vison</i>	0	0	1	0
<i>Nyctereutes procyonoides</i>	1	0	1	0
<i>Ondatra zibethicus</i>	1	0	1	0
<i>Orconectes (Faxonius) limosus</i>	1	1	1	1
<i>Orconectes (Faxonius) rusticus</i>	1	0	1	0
<i>Orconectes (Faxonius) virilis</i>	1	0	1	0
<i>Oxyura jamaicensis</i>	1	1	1	0
<i>Pacifastacus leniusculus</i>	1	0	1	0





<i>Percottus glenii</i>	1	0	1	0
<i>Plotosus lineatus</i>	1	1	1	1
<i>Procamburus clarkii</i>	1	1	1	1
<i>Procamburus fallax</i> f. <i>virginalis</i> (syn. <i>Procamburus virginalis</i>)	1	1	1	1
<i>Procyon lotor</i>	1	1	1	1
<i>Pseudorasbora parva</i>	1	1	1	1
<i>Pterois miles</i>	1	0	0	0
<i>Pycnonotus cafer</i>	1	1	1	1
<i>Pycnonotus jocosus</i>	0	0	1	1
<i>Sciurus carolinensis</i>	1	1	1	0
<i>Sciurus niger</i>	1	0	1	0
<i>Solenopsis geminata</i>	1	1	1	0
<i>Solenopsis invicta</i>	1	1	1	0
<i>Solenopsis richteri</i>	1	0	1	0
<i>Tamias sibiricus</i> (syn. <i>Eutamias sibiricus</i>)	1	1	1	1
<i>Threskiornis aethiopicus</i>	1	1	1	0
<i>Trachemys scripta</i> (syn. <i>Trachemys scripta elegans</i> ; <i>Trachemys scripta scripta</i> ; <i>Trachemys scripta troostir</i>)	1	1	1	1
<i>Wasmannia auropunctata</i>	1	0	1	0
<i>Xenopus laevis</i>	1	1	1	1

- Step 2.** Online retailers and peer-to-peer trade platforms were identified by expert surveyors who speak the respective local languages using a list of search terms (Table 3). The same search terms were used for the identification of international online retail platforms, simply excluding the references to any specific member state. For each identified platform, information was collected on the type of platform (i.e., online retail or peer-to-peer trading platform), the sector it focuses on (i.e., horticulture or pet/aquarium), the country and language it operates in, and the availability of information on IAS in general (i.e., whether there is any content on IAS) as well as on EU and/or platform-specific IAS regulations.

Table 3. Standardised English search terms to identify e-commerce platforms for each of the six MS as well as the internationally operating platforms where the same search terms are utilized without specifying a specific country.

Online retail platform	Peer-to-peer trading platforms
Exotic pets buy <i>MS name</i>	Sell and buy exotic pets <i>MS name</i>
Ornamental fish buy <i>MS name</i>	Sell and buy ornamental fish <i>MS name</i>
Horticulture plants buy <i>MS name</i>	Sell and buy plants <i>MS name</i>

- Step 3.** In 2021, 377 nationally operating platforms were surveyed across the six target MS, as well as 22 internationally operating ones. These 399 platforms included 351 online retail and 48 peer-to-peer trading platforms, and were categorised into 189 horticulture platforms, 159 pet/aquarium platforms, and 51 platforms that offered both





D6.1 Repeatable workflow culturomics

ornamental plants and pets. Numbers of surveyed platforms remained similar in 2024 with 387 separate platforms surveyed in the six MS as well as 10 internationally operating platforms. Here, a total of 397 platforms were included of which 339 were online retail platforms and 58 were peer-to-peer trading platforms. Invasive alien plants and pets were manually searched for in 193 and 213 platforms, respectively. We checked the availability of each IAS on each identified platform using the species' scientific name, the synonyms of scientific names (if relevant), the primary common name(s) in the official language of the MS, and the primary alternative commercial name(s) in the official language of the MS (if relevant). The array of data collected from the search results included information on the identified listings such as the search term used, the species' name under which each IAS is listed, its current or past availability, its life stage and sex, its origin as well as the price. A limitation of this approach is that the physical location of vendors, stock holdings, and shipment origins could not always be verified. Consequently, the presence of a species on a website should not automatically be interpreted as evidence that the species was being traded from within the country where the website was accessible or hosted. The workflow therefore provides an indicator of online availability rather than a direct measure of the geographic origin of trade transactions. Information on the geographic origin of traded species in this case study was unavailable for the majority of records in both survey years (Table 4). For those records where the geographic origin was available, the data indicates trade from both inside and outside the EU across all platform types, sectors, and taxonomic groups with the exceptions of the pet and aquarium sector and the taxonomic grouping "Animals" in 2024, where no trade from outside the EU was documented (Table 4).

Table 4. Number of online trade records by inferred origin category (EU Member States and the UK, outside the EU, mixed origin (only in 2024), or unknown) across platform types, trade sectors, and taxonomic groups in the 2021 and 2024 surveys. Origin categories were derived from information provided by vendors where available. A large proportion of records lacked sufficient information to determine origin and were therefore classified as unknown. The category "Mixed" was introduced in 2024 to account for records indicating multiple countries of origin, including both EU and non-EU countries.

Imported from	Platform type		Sector type			Animal / Plant	
	Online retailer	Peer-to-peer	Horticulture	Pet / Aquarium	Both	Animal	Plant
2021							
EU + UK	22	61	49	19	15	18	65
Outside EU	45	20	19	9	37	11	54
Unknown	496	354	274	244	332	304	546
2024							
EU + UK	80	121	121	48	32	48	153





Outside EU	22	52	55	0	19	0	74
Mixed	0	5	3	0	2	0	5
Unknown	713	384	505	278	312	280	817

- Step 4.** The extracted data was analysed for underlying, temporal trends in the presence of the targeted IAS on e-commerce platforms. First, depending on its presence in e-commerce platforms, each species was classified into “Entry” (i.e. only present in 2024), “Persistent” (i.e. present in 2021 and 2024), and “Exit” (i.e. only present in 2021) (Table 5). Taking into account the species’ status (proposed IAS or listed on the Union list) in 2021 and 2024 as well as whether it was considered in the 2021 analysis, the classification resulted in 10 distinct categories in the presented case study (Table 5). Overall, in 2021, 62 of the 81 surveyed IAS were recorded at least once across the identified e-commerce platforms, while, in 2024, 82 of the 100 included IAS were recorded. Most prominent were those species that were already listed in 2021 and were found in e-commerce in both years. While 12 species were classified as “Exit” consisting of 6 plants (all listed) and 6 animals (4 listed and 2 proposed), only 2 species that were listed and assessed but not found in 2021 (1 plant, 1 animal species) were categorised as “Entry” (Figure 5). Additionally, we explored the relative taxonomic composition (animal/plant) of the IAS found in e-commerce. The ratio of plant and animal species identified in online retailers and peer-to-peer platforms differed between years and countries (Figure 6). Generally, more plant species of Union Concern were present in e-commerce platforms. However, this was not the case in 2024 for online retailers from the Czech Republic and Spain and for internationally operating online retailers. In total, the change of relative plant/animal species proportions shows that online retailers encountered an increase in animal species on their platforms with time, while peer-to-peer trading platforms offered more invasive alien plant species in 2024.

Table 5. Category definitions for classifying the temporal development of a species’ persistence on e-commerce platforms taking into account the species’ status as “listed” or “proposed” for the Union List. Note that the list of categories account for the categories identified in the presented case study and can be adjusted and expanded upon.

Category	Survey_status_2021	e_commerce_2021	Survey_status_2024	e_commerce_2024
Entry - New	Not applicable	Not applicable	Proposed	Present
Entry - Listed	Listed	Absent	Listed	Present
Persistent - Proposed	Proposed	Present	Listed	Present





Persistent - Listed	Listed	Present	Listed	Present
Exit - Proposed	Proposed	Present	Listed	Absent
Exit - Listed	Listed	Present	Listed	Absent
No trade - New	Not applicable	Not applicable	Proposed	Absent
No trade - Proposed	Proposed	Absent	Listed	Absent
No trade - Listed	Listed	Absent	Listed	Absent
Discontinued - Proposed	Proposed	Present	Not applicable	Not applicable

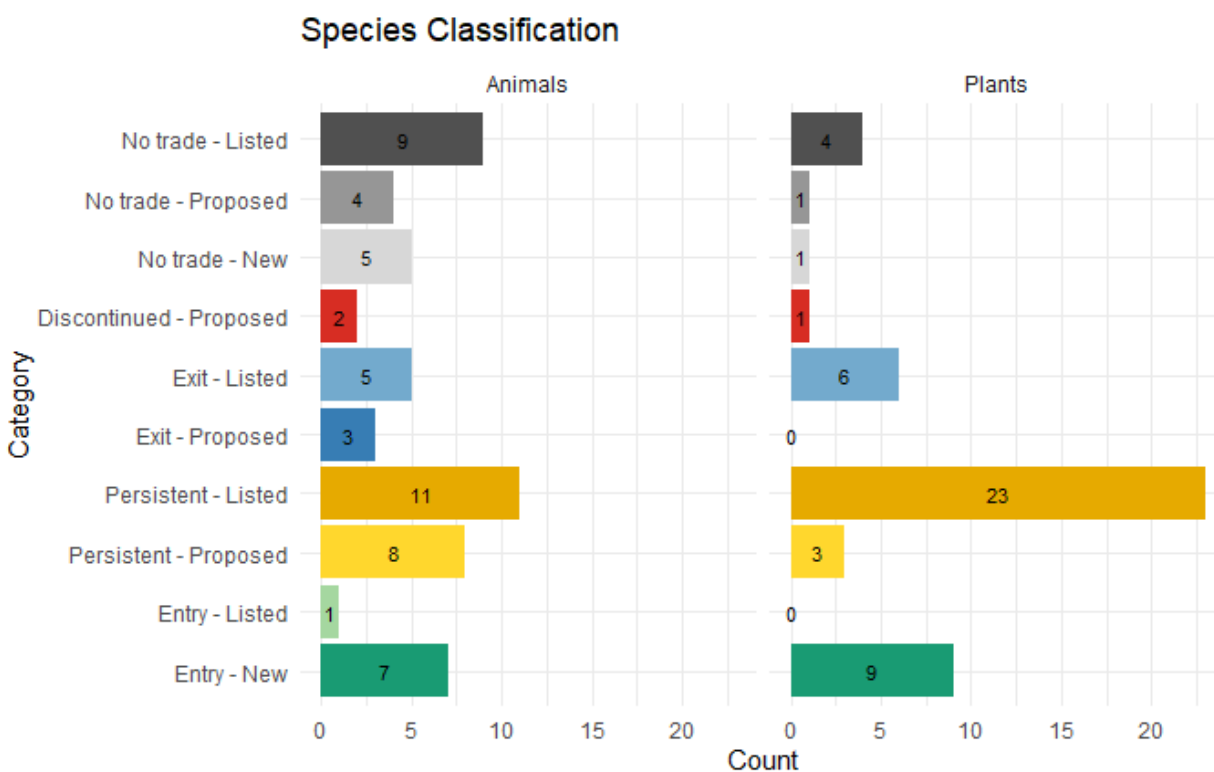


Figure 5. Results of the categorization of the persistence of the presence of IAS of Union concern on e-commerce platforms in 2021 and 2024 based on the categories explained in Table 3.





Species Composition in e-commerce

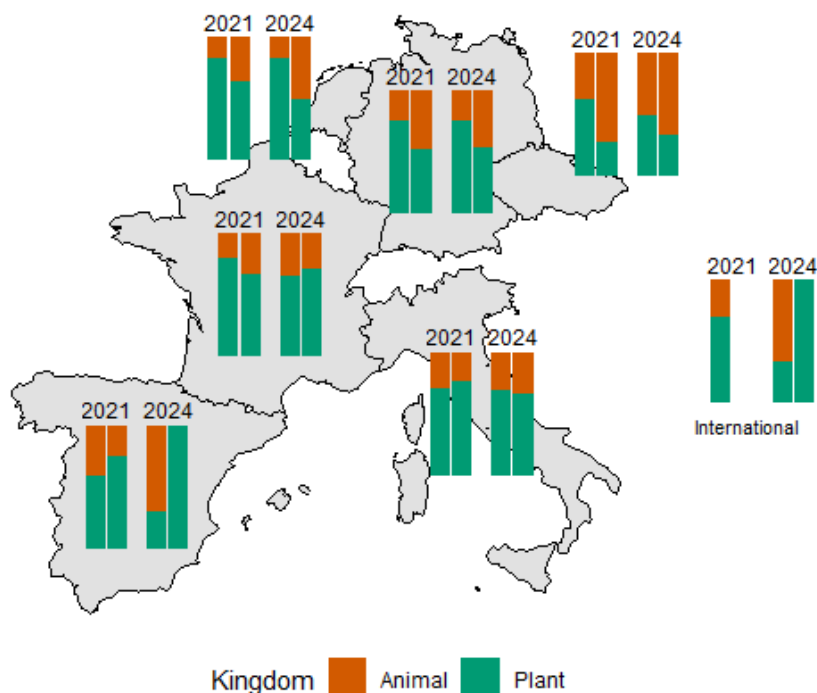


Figure 6. The relative taxonomic composition (plants/animals) of species of Union concern present in e-commerce platforms in 2021 and 2024 for each of the six target EU member states (Czech Republic, France, Germany, Italy, Netherlands, and Spain) as well as a group of internationally operating e-commerce platforms. For each set of two stacked bars per year, the left bar represents “Online retailers” and the right bar “Peer-to-peer trading platforms”.

- **Step 5:** The overall results were used to inform OneSTOP’s Task 6.2.

6. Conclusions

This deliverable develops and documents a repeatable workflow for analysing human interest and behaviours related to IAS using digital data sources. The workflow addresses two important dimensions of biological invasions: public interest in IAS and the continued availability of IAS in e-commerce platforms.

The first sub-workflow, focused on public interest, demonstrates that search engine query data can be used to systematically assess temporal and spatial patterns as well as potential drivers of public interest in IAS across Europe. By combining automated detection of peaks and anomalies with manual contextual analyses, the workflow provides insights into the media’s covered topics associated with increased public interest towards IAS. Such information can support the design of more targeted communication and engagement strategies by identifying which IAS attract the highest public interest, which events trigger spikes in interest, and which





D6.1 Repeatable workflow culturomics

communication channels or actor groups are most influential. For example, information on peaks associated with health concerns, new species detections, or specific media coverage could help practitioners time awareness campaigns, tailor outreach messages to specific audiences, and improve public engagement with IAS management measures.

The second sub-workflow, focused on the availability of IAS in e-commerce platforms, demonstrates that surveys of online retail and peer-to-peer trading platforms can provide valuable information on the persistence and dynamics of IAS trade in e-commerce. It furthermore highlighted the lack of information provided on the geographic origin of where these species are imported from. Despite existing EU regulations, many IAS of Union Concern were still identified on e-commerce platforms during both survey years, underlining the continued importance of monitoring online trade pathways and evaluating compliance with IAS legislation.

A major strength of both sub-workflows is their repeatability, transparency, and scalability. The use of standardised methodologies, templates, and analytical procedures enables comparable data collection across countries, years, and species groups. At the same time, several limitations should be considered. Digital data sources are inherently shaped by platform-specific biases, changing algorithms, language differences, and unequal internet usage across regions and user groups. In addition, some stages of the workflow, particularly the contextual interpretation of public interest peaks, still require manual validation and expert judgement.

Nevertheless, the workflow developed here provides a strong methodological foundation for future activities within WP6 and beyond. Combined with communication and outreach activities, these approaches offer important opportunities to strengthen evidence-based biosecurity monitoring, stakeholder engagement, and policy implementation across Europe.

7. Acknowledgements

We thank the IUCN and DG Environment invasive species teams for providing the data underlying sub-workflow 2.

8. Literature

Alam, M.A. and Hulme, P.E. (2026). Can Google trends be used to estimate the geographic distribution of alien plants in the United States? *Ecological Informatics* 95: 103689.

<https://doi.org/10.1016/j.ecoinf.2026.103689>

Bacher, S., Blackburn, T.M., Essl F, Genovesi, P., Heikkilä, J., Jeschke, J.M., Jones, G., Keller, R., Kenis, M., Kueffer, C., Martinou, A.F., Nentwig, W., Pergl, J., Pyšek, P., Rabitsch, W., Richardson, D.M., Roy, H.E., Saul, W.-C., Scalera, R., Vilà, M., Wilson, J.R.U. and Kumschick, S. (2018). Socio-economic impact classification of alien taxa (SEICAT). *Methods in Ecology and Evolution*, 9(1): 159–168. <https://doi.org/10.1111/2041-210X.12844>

Borchers, H. (2025) pracma: Practical Numerical Math Functions (version 2.4.6). <http://doi.org/10.32614/CRAN.package.pracma>





D6.1 Repeatable workflow culturomics

- Cooper, M. W., Di Minin, E., Hausmann, A., Qin, S., Schwartz, A. J. and Correia, R. A. (2019). Developing a global indicator for Aichi target 1 by merging online data sources to measure biodiversity awareness and engagement. *Biological Conservation* 230: 29–36. <https://doi.org/10.1016/j.biocon.2018.12.004>
- Correia, R., Di Minin, E., Jarić, I., Jepson, P., Ladle, R., Mittermeier, J., Roll, U., Soriano-Redondo, A. and Veríssimo, D. (2019). Inferring public interest from search engine data requires caution. *Frontiers in Ecology and the Environment*, 17(5): 254–255. <https://doi.org/10.1002/fee.2048>
- Dancho, M. and Vaughan, D. (2023). *anomalize: Tidy Anomaly Detection (version 0.3.0)*. <http://doi.org/10.32614/CRAN.package.anomalize>
- Davis, E., Caffrey, J.M., Coughlan, N.E., Dick, J.T.A. and Lucy, F.E. (2018). Communications, outreach and citizen science: Spreading the word about invasive alien species. *Management of Biological Invasions*, 9(4): 515–525. <https://doi.org/10.3391/mbi.2018.9.4.14>
- European Commission: Directorate-General for Environment. (2023). *An introduction to the invasive alien species of Union concern: version 2022*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2779/791022>.
- Gallardo, B., Clavero, M., Sánchez, M.I. and Vilà, M. (2016). Global ecological impacts of invasive species in aquatic ecosystems. *Global Change Biology*, 22(1): 151–163. <https://doi.org/10.1111/gcb.13004>
- Groom, Q., Abraham, L., Adriaens, T., Breugelmans, L., Clarke, D., Di Musciano, M., Dove, S., Estupinan-Suarez, L., Faulkner, K., Fernandez, M., Hendrickx, L., Hui, C., Joly, A., Kumschick, S., Langerhaert, W., Martini, M., Miller, J., Oldoni, D., Pereira, H., Preda, C., Robertson, T., Rocchini, D., Rodrigues, A., Ryckewaert, M., Seebens, H., Sica, Y., Teixeira, H., Trekels, M., Vandaele, T., Wilson, J., Zengeya, T. and Desmet, P. (2025). Creating the vision of rapid, repeatable, reactive data workflows for policy on biodiversity. *Ecological Solutions and Evidence*, 6(3): e70113. <https://doi.org/10.1002/2688-8319.70113>.
- Heiss, R. and Freiling, I. (2026). Addressing social media platforms' influence on academic research. *Humanities and Social Sciences Communications* 13: 192. <https://doi.org/10.1057/s41599-026-06690-6>
- IPBES (2023). *Thematic Assessment Report on Invasive Alien Species and their Control of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Roy, H. E., Pauchard, A., Stoett, P., and Renard Truong, T. (eds.). IPBES secretariat, Bonn, Germany. <https://doi.org/10.5281/zenodo.7430682>
- IUCN. (2018). *Assessing the trade in potential IAS in the European Union*. Technical note prepared by IUCN for the European Commission. TSSR 2018 04
- Jarić, I., Bellard, C., Correia, R.A., Courchamp, F., Douda, K., Essl, F., Jeschke, J.M., Kalinkat, G., Kalous, L., Lennox, R.J., Novoa, A., Proulx, R., Pyšek, P., Soriano-Redondo, A., Souza, A.T., Vardi, R., Veríssimo, D. and Roll, U. (2021). Invasion Culturomics and iEcology. *Conservation Biology*, 35(2): 447–451. <https://doi.org/10.1111/cobi.13707>
- Kumschick, S., Bacher, S., Bertolino, S., Blackburn, T.M., Evans, T., Roy, H.E. and Smith, K. (2020) Appropriate uses of EICAT protocol, data and classifications. *NeoBiota*, 62: 193–212. <https://doi.org/10.3897/neobiota.62.51574>





D6.1 Repeatable workflow culturomics

- Ladle, R.J., Correia, R.A., Do, Y., Joo, G.-J., Malhado, A.C.M., Proulx, R., Roberge, J.-M. and Jepson, P. (2016). Conservation culturomics. *Frontiers in Ecology and the Environment*, 14(5): 269–275. <https://doi.org/10.1002/fee.1260>
- Neumann, K., Mason, S., Farkas, K., Santaularia, J., Ahern, J. and Riddell, C. (2023). Harnessing Google Health Trends Data for Epidemiologic Research. *American Journal of Epidemiology*, 192(3): 430–437. <https://doi.org/10.1093/aje/kwac171>
- Nghiem, L.T.P., Papworth, S.K., Lim, F.K.S. and Carrasco, L.R. (2016). Analysis of the Capacity of Google Trends to Measure Interest in Conservation Topics and the Role of Online News. *PLoS ONE*, 11(3): e0152802. <https://doi.org/10.1371/journal.pone.0152802>
- Novoa, A., Jarić, I., Pipek, P. and Pyšek, P. (2025). Culturomics and iEcology provide novel opportunities to study human and social dimensions of alien species introductions. *Trends in Ecology & Evolution*, 40(1): 18–26. <https://doi.org/10.1016/j.tree.2024.08.012>
- Novoa, A., Dehnen-Schmutz, K., Fried, J. and Vimercati, G. (2017a). Does public awareness increase support for invasive species management? Promising evidence across taxa and landscape types. *Biological Invasions*, 19(12): 3691–3705. <https://doi.org/10.1007/s10530-017-1592-0>
- Novoa, A., Le Roux, J.J., Richardson, D.M. and Wilson, J.R.U. (2017b). Level of environmental threat posed by horticultural trade in Cactaceae. *Conservation Biology*, 31(5): 1066–1075. <https://doi.org/10.1111/cobi.12892>
- Pocock, M.J.O., Adriaens, T., Bertolino, S., Eschen, R., Essl, F., Hulme, P.E., Jeschke, J.M., Roy, H.E., Teixeira, H. and De Groot, M. (2024). Citizen science is a vital partnership for invasive alien species management and research. *iScience*, 27(1): 108623. <https://doi.org/10.1016/j.isci.2023.108623>
- Pyšek, P., Jarošík, V., Hulme, P.E., Pergl, J., Hejda, M., Schaffner, U. and Vilà, M. (2012). A global assessment of invasive plant impacts on resident species, communities and ecosystems: the interaction of impact measures, invading species' traits and environment. *Global Change Biology*, 18(5): 1725–1737. <https://doi.org/10.1111/j.1365-2486.2011.02636.x>
- Reiter, T., Brooks, P.T., Irber, L., Joslin, S.E.K., Reid, C.M., Scott, C., Brown, C.T. and Pierce-Ward, T.N. (2021). Streamlining data-intensive biology with workflow systems. *GigaScience*, 10(1): g1aa140. <https://doi.org/10.1093/gigascience/g1aa140>
- Shackleton, R.T., Shackleton, C.M. and Kull, C.A. (2019a). The role of invasive alien species in shaping local livelihoods and human well-being: A review. *Journal of Environmental Management*, 229: 145–157. <https://doi.org/10.1016/j.jenvman.2018.05.007>
- Shackleton, R.T., Larson, B.M.H., Novoa, A., Richardson, D.M. and Kull, C.A. (2019b). The human and social dimensions of invasion science and management. *Journal of Environmental Management*, 229: 1–9. <https://doi.org/10.1016/j.jenvman.2018.08.041>
- Scharkow, M. and Vogelgesang, J. (2011). Measuring the Public Agenda using Search Engine Queries. *International Journal of Public Opinion Research*, 23(1): 104–113. <https://doi.org/10.1093/ijpor/edq048>
- van Kleunen, M., Essl, F., Pergl, J., Brundu, G., Carboni, M., Dullinger, S., Early, R., González-Moreno, P., Groom, Q.J., Hulme, P.E., Kueffer, C., Kühn, I., Máguas, C., Maurel, N., Novoa, A., Parepa, M., Pyšek, P., Seebens, H., Tanner, R., Touza, J., Verbrugge, L., Weber, E., Dawson, W., Kreft, H., Weigelt, P., Winter, M., Klöner, G., Talluto, L. and Dehnen-Schmutz, K.





D6.1 Repeatable workflow culturomics

(2018). The changing role of ornamental horticulture in alien plant invasions. *Biological Reviews*, 93(3): 1421–1437. <https://doi.org/10.1111/brv.12402>

Vardi, R., Mittermeier, J. C. and Roll, U. (2021). Combining culturomics sources to uncover trends in popularity and seasonal interest in plants. *Conservation Biology* 35(2): 460–471. <https://doi.org/10.1111/cobi.13705>

Vilà, M., Espinar, J.L., Hejda, M., Hulme, P.E., Jarošík, V., Maron, J.L., Pergl, J., Schaffner, U., Sun, Y. and Pyšek, P. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology Letters*, 14(7): 702–708. <https://doi.org/10.1111/j.1461-0248.2011.01628.x>

