

Reglement für die Datenfreigabe durch Mediapulse

Mediapulse Online Content Data

Sie wollen wissen, welche Kriterien für eine Publikation erfüllt sein sollen?

Mediapulse Online Data werden vor der Publikation einer Qualitätsprüfung unterzogen, die nach festgelegten Kriterien durchgeführt wird. Im folgenden Dokument werden der Qualitäts-Prüfungs-Prozess, die zugrundeliegenden Kriterien und Schwellenwerte, sowie die Massnahmen für die Publikation beschrieben.

1. Das QA Monitoring führt zur Identifikation von Messfehlern

Die Ursachen für fehlerhafte Daten können einerseits bei der Implementierung des Messtags entstehen, andererseits ist es denkbar, dass bei der Datenverarbeitung (Dataprocessing) Probleme auftauchen und gewisse Tage nicht gemessen werden können.

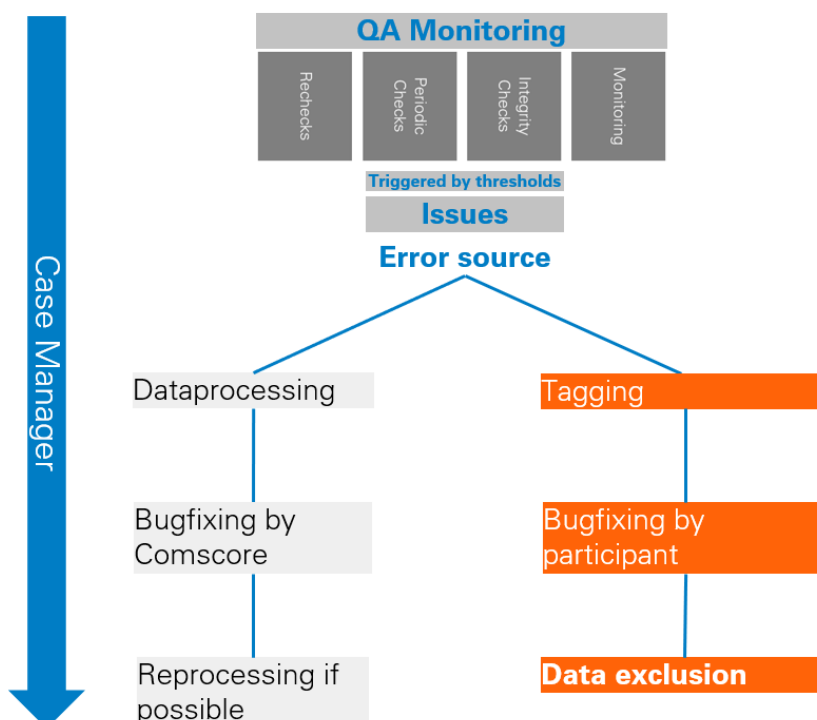


ABBILDUNG 1: QA MONITORING, FEHLERQUELLEN UND DATENAUSSCHLUSS

2. Fehlerhafte Daten können zu einem Datenausschluss führen

Die Daten eines Brands werden auf Tagesbasis aus der Publikation ausgeschlossen, wenn sie die Qualitätsanforderungen gemäss Vorgaben des QA Monitorings nicht erfüllen.

- Für einen Datenausschluss muss (a) ein eindeutiger Fehler (Error Source) identifiziert sein, der (b) eine publizierte Metrik übermässig (definiert durch Schwellenwerte im QA Monitoring) beeinflusst.
- Daten werden nur beim Nachweis von Bot-Traffic statistisch korrigiert. Im Fall von Fehlern basierend auf einer falschen Messung werden die Daten nicht statistisch korrigiert.
- Sind 20 oder mehr Tage eines Monats korrekt gemessen, hat der Teilnehmer die Möglichkeit an der Monatspublikation trotz ausgeschlossenen Tagen teilzunehmen. In diesem Fall werden ausschliesslich Metriken auf Basis von Durchschnittswerten in der Publikation ausgewiesen.
- Sind weniger als 20 korrekte Messtage vorhanden, kann für den spezifischen Brand im betroffenen Monat keine Publikation stattfinden.

3. Im QA Monitoring werden Daten einer Qualitätsprüfung unterzogen

Die Qualitätsprüfung erfolgt über das QA Monitoring, welches kontinuierlich unter klar definierten Standards (Schwellenwerte, Toleranzwerte) stattfindet. Die folgenden drei Elemente der Qualitätsprüfung sind relevant für die Publikation, d.h. sie können zu Datenausschluss oder signifikanten Anpassungen in der Publikation (z.B. Ausschluss invalider Brands) führen.

3.1. *Schwellenwerte lösen Abklärungen aus und können zu Datenausschluss führen*

3.1.1 *Identifikation:*

Über Schwellenwerte werden im QA Monitoring Auffälligkeiten im Traffic-Verlauf auf Brand und Asset-Level identifiziert:

- Modus: Werktätlich; am Montag für Fr, Sa und So
- Verantwortlichkeit: Case Manager bei Kantar
- Gegenstand: Tägliche Visits, IDs, Durations auf Basis Device und Betriebssystem werden auf Auffälligkeiten geprüft.

3.1.2 Observation:

Sobald der beobachtete Wert eines Brands oder Assets während mehr als zwei aufeinanderfolgenden Tagen mehr als zwei Standardabweichungen über oder unter dem Mittelwert der letzten 28 Tage liegt, werden erste Abklärungen ausgelöst:

Abklärungen zur

- a) Identifikation einer inhaltlichen Ursache (falls nötig, wird Kontakt mit den Verantwortlichen aufgenommen)
- b) Identifikation eines technischen Messfehlers inkl. Unterscheidung zw. Fehler bei Comscore oder im Tagging sowie sofortige Einleitung von Massnahmen zur Behebung (Reprocessing oder Bugfixing)
- c) Identifikation von Indizien, die auf nicht-nutzerinduzierten Traffic (Bot) hinweisen.

3.1.3 Massnahme Exklusion:

- a) Falls ein Fehler in der Messung seitens Tagging oder ein Fehler in den Messdaten identifiziert wurde (siehe 3.1.2 a) oder b)) **und**
- b) der Messwert des Brands mehr als 2 Tage lang über oder unter der doppelten Standardabweichung liegt **und**
- c) keine inhaltlichen Gründe den Ausreisser erklären, sind die Kriterien für den Ausschluss eines spezifischen Brands aus der Publikation erfüllt.

Der Teilnehmer wird über den Ausschluss einzelner Tage¹ oder des gesamten Monats² informiert. In eindeutigen Fällen (eindeutig identifizierter Fehler) kann auch bereits nach einem Tag übermässiger Abweichung ein Ausschluss einzelner Messtage erfolgen.

Falls ein Fehler in der Messung seitens Comscore vorliegt, wird alles darangesetzt, via Reprocessing den Fehler zu korrigieren.

Kriterien		Observation	Ausschluss
Schwellenwert		Standardabweichung (+/- 2) mehr als 2 Tage überschritten	Mehr als 3 Tage Standardabweichung (+/- 2) überschritten
Technischer Messfehler	Tagging	Fehlersuche/ Bugfixing	Fehler/ Issue nachgewiesen
	Comscore	Fehlersuche/ Reprocessing	Kein Fehler/ Issue seitens Comscore gefunden
Inhaltliche Ursache		Inhaltliche Abklärungen	Keine inhaltliche Ursache gefunden

ABBILDUNG 2: KRITERIENMATRIX DATENAUSSCHLUSS

¹ Falls trotzdem mind. 20 korrekte Messtage des betroffenen Monats vorhanden sind

² Falls weniger als 20 korrekte Messtage des betroffenen Monats vorhanden sind

3.1.4 Massnahme Korrektur ausschliesslich bei Bot-Traffic:

Im Falle von eindeutig identifiziertem Bot-Traffic wird der Publisher kontaktiert und eine Daten-Korrektur vorgenommen. Es handelt sich um eindeutig identifizierten Bot-Traffic wenn

- a) der übermässige Traffic (Schwellenwerte siehe Observation) auf einzelne User Agents oder IPs zurückzuführen ist und
- b) der Publisher diese Beobachtung bis Ende Monat nicht widerlegen kann, resp. keine inhaltlichen Gründe aufführen kann, die auf nutzerinduzierten Traffic hinweisen

Die Korrektur wird folgendermassen vorgenommen:

Die Korrektur basiert auf den letzten 4 gleichen Wochentagen. (Mittelwerts-Ersetzung: Ausfall Montag wird ersetzt mit dem Mittelwert aus den letzten 4 Montagen)

Korrigiert werden die Daten auf der granularsten Stufe und jeweils in die übergeordnete Aggregation übernommen. D.h. der Bot-Traffic wird pro Brand / Asset / OS und Geolokalisation korrigiert. Die Korrektur wird in die Brand- und Netzwerksicht zur Publikation übernommen.

Im Gegensatz zum Bot-Traffic werden Qualitätsmängel auf Grund von Messfehlern nicht korrigiert. Die Massnahmen zum Umgang mit Bots orientieren sich an den uns bekannten Vorkommnissen, es ist nicht ausgeschlossen, dass neue Erkenntnisse neue Massnahmen erfordern.

3.2. Schwellenwerte zur Kontrolle Asset und Brandbezeichnungen

Schwellenwerte werden auch eingesetzt, um Asset und Brandbezeichnungen zu kontrollieren:

- Modus: Werktätlich; am Montag für Fr, Sa und So
- Verantwortlichkeit: Case Manager bei Kantar
- Gegenstand: Bei der Validierung werden Brand- und Asset-Namen mit dem Intake-Formular abgeglichen, invalide Brands und Assets werden von der Publikation ausgeschlossen

Asset	Brand	Massnahmen
VALID	VALID	Messdaten sind valide
VALID	FALSE	Ausschluss aus den Messdaten
FALSE	VALID	Sobald die Summe der nicht gültigen Assets bei einer beliebigen Metrik auf tagesbasis 5 % des Brandtotals überschreitet, wird der Brand ausgeschlossen. Liegt der Traffic, der auf falsche Asset-Bezeichnungen beruht, unter 5%, gilt der Brand noch als validiert.
FALSE	FALSE	Ausschluss aus den Messdaten

ABBILDUNG 3: VALIDE VS. INVALIDE ASSET- UND BRANDBEZEICHNUNGEN

Falls das Total der nicht validierten Assets über 2% des Brand Totals liegt, wird der Teilnehmer von Mediapulse kontaktiert.

3.3. Toleranzwert zum Anteil an Fremd-Domains bei Web-Assets

Es gibt Einschränkungen bei der Bündelung von Domains. Dafür wird regelmässig der Anteil an Fremd-Domains geprüft:

- Modus: monatlich
 - Verantwortlichkeit: Case Manager bei Kantar
 - Gegenstand: Die regelmässige Kontrolle zur Begrenzung von Fremd Domains kann zu Korrekturen bei der Domain-Bündelung führen
1. Bündelung von Assets zu Brands: Das Hauptangebot* (ns_site) muss mehr als 50% der gesamten Webseitenbesuche eines Brands ausmachen. Damit wird eine künstliche Inflation der Daten durch Bündelung mehrerer Web-Angebote begrenzt.
 2. Nutzung von mehreren Domains bei Assets: Die Top-Level-Domain (oder Bündel davon) mit dem gleichen Namen wie das Asset muss mehr als 50% der Visits der beobachteten Domains in der Messung ausmachen. Damit wird der Einfluss der Tag-Einbindung auf fremde Domains begrenzt.

4. Kundeninformationsfrist vor der Datenfreigabe

Der Teilnehmer erhält bei publikationsrelevanten Themen, welche nach einem Entscheid des Teilnehmers verlangen, bis spätestens 5 Arbeitstage vor Datenfreigabe die dafür notwendigen Informationen.

Anpassungen:

Version	Date/ Name	Kommentar
Version 1	Juni 2021/ Isabelle Waser	
Version 2	Sept 2021/ Isabelle Waser	Neu: Kapitel 3.1.4: Korrektur bei Bot-Traffic Neu: Kapitel 4: Kundeninformationsfrist vor Datenfreigabe Korrektur: Kapitel 3.2 Abbildung 3 ergänzt mit "tagesbasiert"