

## D4.2 Finales Konzept und Schemata für die Schadensharmonisierung

28.02.2022

VERSION:	Version 2.0
STATUS:	draft / <u>final</u>
DATUM:	28.02.2022
ERSTELLT:	Judith Köberl
AUTOREN:	Judith Köberl, Katharina Enigl, Sebastian Lehner
FREIGEgeben DURCH:	Matthias Themeßl

## KURZZUSAMMENFASSUNG

---

In Arbeitspaket (AP) 4 werden die Grundlagen und Konzepte zur Transformation und Harmonisierung der aus heterogenen Quellen stammenden Ereignis- und Schadendaten erstellt und in einem iterativen Prozess getestet, angepasst und angewendet. Die Anwendung und damit der Funktionsnachweis der erstellten Konzepte erfolgt auf Ereignis- und Schadendaten aus den zwei Bundesländern Niederösterreich und Steiermark unter Einbezug der drei Gefahrenkategorien Wind/Sturm, Hochwasser/Überflutung und Massenbewegungen.

Das vorliegende Deliverable beschreibt den vollständigen, im Rahmen von CESARE durchgeführten Harmonisierungsprozess, wobei der ex-post bzw. retrospektiven Harmonisierung – wie sie in CESARE erfolgt – gewisse Grenzen gesetzt sind. Der retrospektive Datenharmonisierungsprozess in CESARE erfolgt in fünf Schritten:

- 1.) Definition der gewünschten Zielschemata und -formate für jene Merkmale bzw. Variablen, die harmonisiert werden sollen: Datums-/Zeitangaben, Verortung, Art der Naturgefahr (Hazard), Schadenkategorie bzw. Art des geschädigten Elements, Schadausmaß, Objektbesitzer:in, Schadenträger:in, Datenquelle, Ereignis-ID, Komposit-ID.
- 2.) Analyse des jeweils betrachteten Originaldatensatzes in Bezug auf relevante Informationen, Harmonisierungserfordernisse und Harmonisierungspotential. Das Harmonisierungspotential wird dabei mithilfe einer fünfstufigen Skala bewertet.
- 3.) Erstellung möglichst allgemeiner und (teil-)automatisierter Regeln und Prozesse, um die Quellschemata in die harmonisierten Zielschemata zu überführen.
- 4.) Anwendung der in Schritt 3 erstellten Regeln und Prozesse sowie des zur Ereigniszusammenführung bzw. Ereignis- und Komposit-ID-Vergabe entwickelten Algorithmus auf den jeweils betrachteten Original-Datensatz.
- 5.) Einspielung des harmonisierten Datensatzes in den CESARE Demonstrator.

Um im Zielschema Unsicherheiten bzw. den Grad der Zuverlässigkeit einzelner Angaben darzustellen, wird auf den MAXO-Code zurückgegriffen, der beispielsweise im WLK-Datensatz und im HFWDB-Datensatz zur Anwendung kommt. Neben der Harmonisierung selbst findet auch eine Bewertung der Originaldaten und des Harmonisierungsprozesses statt. Dies erfolgt anhand der folgenden Kategorien: Informationsgehalt/Detailgrad, Unsicherheiten/Unschärfen, Angaben zur Messgenauigkeit, Harmonisierungspotential und Informationsverlust durch Harmonisierung.

Auch wenn die im Rahmen von CESARE verfügbaren Datensätze noch bei weitem nicht das gesamte Schadvolumen für die drei hauptsächlich betrachteten Naturgefahren (Hochwasser, Sturm, Massenbewegungen) erfassen, zeigt sich, dass mit ihrer Hilfe Konzepte erstellt und erprobt werden können, wie man Daten aus unterschiedlichen Quellen transformieren und zusammenführen kann, sodass eine gemeinsame Auswertung – unter Berück-

sichtigung verbleibender Unsicherheiten und Unschärfen – möglich ist. Zudem lassen sich die Informationsgewinne und der Mehrwert der Zusammenführung und gemeinsamen Nutzung unterschiedlicher Daten illustrieren.

# INHALT

---

<b>KURZZUSAMMENFASSUNG</b>	<b>2</b>
<b>INHALT</b>	<b>4</b>
<b>1. EINLEITUNG</b>	<b>5</b>
<b>2. ALLGEMEINE VORGEHENSWEISE</b>	<b>6</b>
<b>3. DATENHARMONISIERUNG IN CESARE</b>	<b>8</b>
<b>4. VERBLEIBENDE HERAUSFORDERUNGEN</b>	<b>47</b>
<b>5. QUALITÄT DER DATEN UND DES HARMONISIERUNGSPROZESSES</b>	<b>42</b>
<b>6. REFERENZEN</b>	<b>49</b>
<b>A. ANHANG</b>	<b>49</b>

# 1. EINLEITUNG

---

In Arbeitspaket (AP) 4 werden die Grundlagen und Konzepte zur Transformation und Harmonisierung der aus heterogenen Quellen stammenden Ereignis- und Schadendaten erstellt und in einem iterativen Prozess getestet, angepasst und angewendet. Die Anwendung – und damit der Funktionsnachweis der erstellten Konzepte – erfolgt auf Ereignis- und Schadendaten aus den zwei Bundesländern Niederösterreich und Steiermark unter Einbezug der drei Gefahrenkategorien Wind/Sturm, Hochwasser/Überflutung und Massenbewegungen.

Verschiedenste Institutionen erfassen und sammeln in Österreich derzeit Ereignis- und/oder Schadendaten zu Naturgefahren. Dies erfolgt aus unterschiedlichen Gründen und Motivationen heraus, wobei nicht immer die Ereignis- oder Schadendatensammlung an sich im Vordergrund steht. Sie ergibt sich teilweise als „Nebenprodukt“, wie etwa im Falle der Beihilfenabwicklung von Katastrophenschäden durch die Landesverwaltung. Methode und Umfang der Datensammlung sind hier insbesondere auf den primären Zweck ausgerichtet: die Abwicklung von Beihilfenauszahlungen (Köberl et al., 2018). So kommt es, dass sich die unterschiedlichen Ereignis- und Schadendatensammlungen in Erhebungszweck, -methodik, -umfang und -standard stark unterscheiden können.

Es ist weder realistisch noch uneingeschränkt sinnvoll, dass Institutionen, die derzeit Ereignis- und Schadendaten zu Naturgefahren sammeln, ihre originären Datenmodelle aufgeben und ihre interne Datenhaltung an die CESARE Datenmodelle anpassen. Allerdings erschwert die Heterogenität zwischen den unterschiedlichen vorhandenen Datensätzen die Wiederverwendung der erhobenen Daten durch andere Organisationen sowie die Zusammenführung und gemeinsame Nutzung mit anderen Daten, aus der ein Mehrwert an Information entstehen kann. AP 4 widmet sich daher der Harmonisierung und Standardisierung von aus unterschiedlichen Quellen stammenden Ereignis- und Schadendaten, um die Heterogenität ex-post zu reduzieren und die Vergleichbarkeit zu erhöhen. Die entwickelten Konzepte, Zielschemata und Standards sind in zwei Berichten festgehalten:

- **Deliverable 4.1 (D4.1): Zielschemata für die Transformation von Events**

In D4.1 werden (i) die für den Demonstrator zur Verfügung gestellten Datensätze beschrieben, (ii) die entwickelten Zielschemata für die Bereiche „Gefahren/Hazards“ sowie „Schadensindikatoren und betroffene Elemente“ vorgestellt und (iii) die Anwendung der entwickelten Zielschemata auf die verfügbaren Datensätze demonstriert. Dabei handelt es sich bereits um einen Teil des Harmonisierungsprozesses, der im vorliegenden Dokument ausführlicher beschrieben wird.
- **Deliverable 4.2 (D4.2): Finales Konzept und Schemata für die Schadensharmonisierung**

Der vorliegende Bericht beschreibt den vollständigen, im Rahmen von CESARE durchgeführten Harmonisierungsprozess, wobei vorwegzunehmen ist, dass der ex-post Harmonisierung und Standardisierung gewisse Grenzen gesetzt sind.

Informationen aus unterschiedlichen Datenquellen können sich komplementär ergänzen, überschneiden oder widersprechen. Ziel von D4.2 ist die Erarbeitung von Kriterien und Regeln für das Zusammenführen und Kombinieren quantitativer Aussagen aus unterschiedlichen, heterogenen Quellen, die im Hinblick auf automatisierte Auswertungen bzw. Aggregationen herangezogen werden können. Diesbezüglich zu klärende Fragen umfassen

beispielsweise den Umgang mit unterschiedlichen Schadenangaben, mit unterschiedlichen Schadensdeckelungen oder mit Unterschieden in der Schadenkategorisierung bzw. -definition in den Primärquellen. Des Weiteren gilt es, die Datenqualität der Primärquellen, aber auch die Qualität der harmonisierten Daten zu beurteilen.

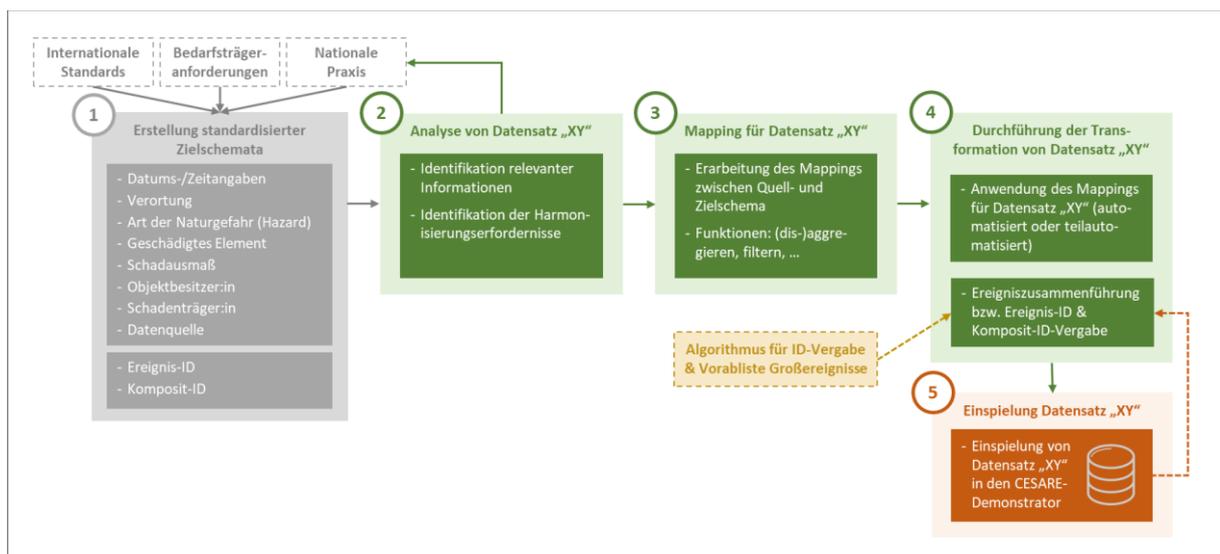
## 2. ALLGEMEINE VORGEHENSWEISE

Innerhalb von CESARE geht es um eine retrospektive Datenharmonisierung. Durch Harmonisierung wird versucht, verschiedene Arten, Ebenen und Quellen von Daten, die die Messung derselben latenten Konstrukte darstellen, so zusammenzuführen, dass sie möglichst kompatibel und vergleichbar gemacht werden können. Daten können dabei neu codiert, transformiert oder mit zusätzlichen Informationen kombiniert werden, um eine Harmonisierung zu erreichen. Die retrospektive Datenharmonisierung umfasst grob gesprochen drei Schritte (vgl. MRC, 2017):

- 1.) **Definition der Zielvariable(n)**, d. h. des gewünschten gemeinsamen Formats, das durch Harmonisierung von den verfügbaren Original- bzw. Rohdaten hergeleitet werden soll.
- 2.) **Bewertung des Harmonisierungspotentials**, u. a. auf Basis von Metainformationen zu den Originaldaten. Nicht immer muss das Potential zur Harmonisierung gegeben sein.
- 3.) **Ableitung von Daten gemeinsamen Formats**, beispielsweise durch (i) die Konvertierung der Maßeinheit, (ii) die Verwendung von Validierungsdaten oder (iii) das Hinunterbrechen auf den kleinsten gemeinsamen Nenner, was auch die Verwendung von Annahmen beinhalten oder einen Verlust von Detailliertheit mit sich bringen kann.

Rund um diese drei groben Schritte der Datenharmonisierung wird der CESARE-Datenharmonisierungsprozess aufgebaut. Abbildung 2-1 gibt einen Überblick über das Ablaufschema dieses Harmonisierungs- und Standardisierungsprozesses. Eine detaillierte Beschreibung folgt in Kapitel 3.

Abbildung 2-1: Ablaufschema der Datenharmonisierung und -standardisierung



Quelle: eigene Darstellung

Der **erste Schritt** besteht aus der Definition der gewünschten Zielschemata und -formate für jene Merkmale bzw. Variablen, die im Zuge von CESARE harmonisiert werden sollen. Die Definition der Zielschemata und -formate erfolgt in CESARE nach Möglichkeit auf Basis (i) internationaler Standards und Empfehlungen, (ii) nationaler Standards und Praktiken (abgeleitet aus den verfügbaren Datensätzen) sowie (iii) unter Berücksichtigung der Bedarfsträgeranforderungen. Folgende Merkmale bzw. Variablen werden im Zuge von CESARE harmonisiert:

- Datums-/Zeitangaben
- Verortung
- Art der Naturgefahr (Hazard)
- Schadenkategorie bzw. Art des geschädigten Elements
- Schadausmaß
- Objektbesitzer:in (bei Sachschäden)
- Schadenträger:in
- Datenquelle
- Ereignis-ID
- Komposit-ID

Im **zweiten Schritt** erfolgt die Analyse des jeweils betrachteten Originaldatensatzes in Bezug auf relevante Informationen für den CESARE Demonstrator, aber auch in Bezug auf Harmonisierungserfordernisse und -potential. Je Zielvariable wird das Quellschema der Originaldaten mit dem Zielschema der Zielvariable abgeglichen und das Harmonisierungspotential bewertet. Dabei kommen, sofern verfügbar, auch Metainformationen zu den Originaldaten zum Einsatz. Nicht immer ist das Potential zur Harmonisierung gegeben. Ist das Harmonisierungspotential niedrig oder nicht gegeben, gilt es, zwischen der Reduktion des Harmonisierungsgrades bei Aufnahme des Datensatzes und dem Fehlen wichtiger Informationen bei Weglassen des Datensatzes abzuwägen.

Bei den definierten Zielschemata handelt es sich im Übrigen um keine komplett starren Konstrukte. Legt die Analyse eines neuen Originaldatensatzes eine sinnvolle Erweiterung oder Anpassung eines Zielschemas nahe, sind entsprechende Abänderungen möglich bzw. angedacht.

Im **dritten Schritt** werden – basierend auf dem Abgleich zwischen Quell- und Zielschemata – möglichst allgemeine und (teil-)automatisierte Regeln und Prozesse erstellt, um die Quellschemata in die harmonisierten Zielschemata zu überführen. Diese Regeln und Prozesse können sich je nach Originaldatensatz und je nach Zielvariable deutlich unterscheiden.

Der **vierte Schritt** umfasst die Anwendung der im dritten Schritt erstellten Regeln und Prozesse. Des Weiteren wird der in Kapitel 3 detailliert beschriebene Algorithmus zur Ereigniszusammenführung bzw. Ereignis- und Komposit-ID-Vergabe auf den betrachteten Datensatz angewendet, wobei neben einer vorab definierten Liste an Großereignissen auch etwaige bereits in den CESARE-Demonstrator eingespielte Ereignisse berücksichtigt – und falls notwendig redefiniert – werden.

Nach abgeschlossener Harmonisierung erfolgt im **fünften Schritt** die Einspielung des Datensatzes in den CESARE Demonstrator.

### 3. DATENHARMONISIERUNG IN CESARE

---

Für jedes CESARE-Merkmal folgt eine Beschreibung (i) des Zielschemas, (ii) des Harmonisierungsbedarfs und -potentials der einzelnen Originaldatenquellen in Bezug auf das betrachtete Merkmal, (iii) der Bearbeitungsschritte zur Überführung des Quellschemas in das Zielschema und (iv) etwaiger verbleibender Herausforderungen. Die qualitative Bewertung des Harmonisierungspotentials erfolgt dabei anhand einer fünfstufigen Skala:

- niedrig
- mittel-niedrig
- mittel
- mittel-hoch
- hoch

#### DATUMS-/ZEITANGABEN

##### Definition des CESARE Zielschemas

Datums- und Zeitangaben zu Ereignissen und Schäden unterscheiden sich zwischen und zum Teil auch innerhalb der zur Verfügung stehenden Originaldaten<sup>1</sup> u. a. in Bezug auf das Format, den Detailgrad, den Grad der Zuverlässigkeit und der Kennzeichnung des Grades der Zuverlässigkeit. In CESARE zielen wir in Bezug auf den Detailgrad auf die Tagesebene ab. Etwaige im Original vorhandene Informationen zur Uhrzeit eines Ereignisses oder Schadeintritts werden daher vorläufig nicht ein einheitliches Zielschema überführt. Was Datumsangaben betrifft, wird das Zielschema derart gewählt, dass es möglichst flexibel auf die unterschiedlichsten Formate und Detailgrade in den Originaldaten eingehen und etwaige Informationen zum Grad der Zuverlässigkeit der Datumsangabe erfassen kann. Das CESARE Zielschema für das Beginn- und Enddatum von Ereignissen und Schäden sieht daher die Trennung der einzelnen Datumskomponenten (Jahr, Monat, Tag) in jeweils separate Merkmale vor. Der Grad der Zuverlässigkeit der einzelnen Datumskomponenten wird anhand des MAXO-Codes (M= Messwert/Feststellung, A = Annahme/Schätzung, X = unklar/noch zu erheben, O = nicht bestimmbar), der beispielsweise im WLK-Datensatz und im HFWDB-Datensatz zur Anwendung kommt, ausgedrückt. Neben Beginn- und Enddatum stellt auch die Dauer, gemessen in Tagen, eine CESARE-Zielvariable dar. Sie kann (i) genau der Differenz zwischen Beginn- und Enddatum entsprechen, (ii) kleiner als die Differenz zwischen Beginn- und Enddatum sein, falls es Unsicherheiten zum genauen Beginn oder Ende gibt und nur bekannt ist, dass ein Ereignis und/oder Schaden bestimmter Dauer irgendwann zwischen dem angeführten Beginn- und Enddatum aufgetreten ist oder

---

<sup>1</sup> Für eine Beschreibung der zur Verfügung stehenden Originaldaten siehe auch CESARE-Deliverable D4.1.

(ii) leer sein, wenn ein Ereignis oder Schadeintritt unbekannter Dauer zwischen Beginn- und Enddatum stattgefunden hat. Folgende Merkmale werden somit im CESARE Zielschema in Bezug auf den Ereignis- bzw. Schadenzeitraum erfasst:

- Beginn\_Jahr (integer)
- Beginn\_Jahr\_MAXO (enumeration: {M, A, X, O})
- Beginn\_Monat (integer: {1-12})
- Beginn\_Monat\_MAXO (enumeration: {M, A, X, O})
- Beginn\_Tag (integer: {1-31})
- Beginn\_Tag\_MAXO (enumeration: {M, A, X, O})
- Ende\_Jahr (integer)
- Ende\_Jahr\_MAXO (enumeration: {M, A, X, O})
- Ende\_Monat (integer: {1-12})
- Ende\_Monat\_MAXO (enumeration: {M, A, X, O})
- Ende\_Tag (integer: {1-31})
- Ende\_Tag\_MAXO (enumeration: {M, A, X, O})
- Dauer (integer)
- Dauer\_MAXO (enumeration: {M, A, X, O})

## Harmonisierungspotential und Umsetzung je Datensatz

### GEORIOS

In GEORIOS ist das Ereignisdatum als Text gespeichert. Der Detailgrad variiert innerhalb des Datensatzes und reicht vom Jahr des Ereignisses bis zur Uhrzeit. Angaben zur Zuverlässigkeit bzw. etwaige Unsicherheiten in den Zeitangaben werden mit den Codierungen „/“ (für „oder“) und „?“ (für „fraglich“) ausgedrückt. Das Harmonisierungspotential ist **hoch**, allerdings kann die Harmonisierung aufgrund vieler Sonderfälle nur semi-automatisiert erfolgen. Tabelle 3-1 zeigt einige Beispiele, die das Mapping zwischen Quell- und Zielschema illustrieren. Zeitangaben werden, wie erwähnt, nicht ins Zielschema überführt. Bei im Original fehlenden Tages- oder Monatsangaben bleiben die entsprechenden CESARE-Merkmale leer und der zugehörige MAXO-Wert wird auf „O“ (nicht bestimmbar) gesetzt. Im Original durch „/“ und „?“ ausgedrückte Unsicherheiten werden im Zielschema durch ein „A“ (Annahme/Schätzung) im entsprechenden MAXO-Feld gekennzeichnet. Jahreszeitangaben wie „Sommer“, „Winter“, „Herbst“ etc. werden meteorologisch interpretiert, wobei die mit der Bezeichnung einhergehende Unschärfe durch ein „A“ in den MAXO-Werten der Merkmale Beginn\_Monat und Ende\_Monat gekennzeichnet wird.

Tabelle 3-1: Beispiele zum Mapping zwischen Quell- und Zielschema der Datums- und Zeitangaben in GEORIOS

Originalangabe	Beginn			Beginn_MAXO			Ende			Ende_MAXO			Dauer	Dauer_MAXO
	jjjj	m	t	jjjj	m	t	jjjj	m	t	jjjj	m	t		
2011-04-17 (08:06 Uhr)	2011	4	17	M	M	M	2011	4	17	M	M	M	1	M
2011-06-11 (gegen 05:30 Uhr)	2011	6	11	M	M	M	2011	6	11	M	M	M	1	M
2018-12-08 (vormittags)	2018	12	8	M	M	M	2018	12	8	M	M	M	1	M
2006-04-26 00:00:00	2006	4	26	M	M	M	2006	4	26	M	M	M	1	M
2010-08	2010	8		M	M	O	2010	8		M	M	O		O
2005	2005			M	O	O	2005			M	O	O		O
2015-10-15/16	2015	10	15	M	M	A	2015	10	16	M	M	A	1	A
Sommer 2012	2012	6		M	A	O	2012	8		M	A	O		O
2009-06-24 bis 2009-06-26	2009	6	24	M	M	M	2009	6	26	M	M	M	3	M
2014-10-29(?)	2014	10	29	M	M	A	2014	10	29	M	M	A	1	A
2009-06-23 (?)	2009	6	23	A	A	A	2009	6	23	A	A	A	1	A

jjjj= Jahr, m = Monat, t = Tag; Originalangabe: Codierung mit „?“ ohne Leerzeichen bezieht sich auf den Teil der Datumsangabe, bei dem das Fragezeichen steht. Codierung mit „?“ nach dem Datum und mit Leerzeichen bezieht sich auf die gesamte Datumsangabe.

#### WLK

Im WLK-Datensatz wird der Ereigniszeitraum sowie die Zuverlässigkeit der Angaben zum Ereigniszeitraum mithilfe von sieben Merkmalen ausgedrückt:

- ereignis\_datum: vollständiges Ereignisdatum in der Form jjjj/mm/tt, sofern bekannt
- ereignis\_zeitpunkt\_maxo: der sich auf den Ereigniszeitpunkt beziehende MAXO-Wert
- ereignis\_jahr: Jahr des Ereignisses, falls das vollständige Ereignisdatum nicht bekannt ist; zusätzlich zum Jahr des Ereignisses kann auch das Monat (ereignis\_monat) angeführt sein
- ereignis\_monat: Monat des Ereignisses, falls das vollständige Ereignisdatum nicht bekannt ist
- ereignis\_dauer: Dauer des Ereignisses in Tagen, Stunden, Minuten, Sekunden, sofern bekannt
- ereignis\_dauer\_text: Textangabe zur Ereignisdauer
- ereignis\_dauer\_maxo: der sich auf die Ereignisdauer beziehende MAXO-Wert

Das Harmonisierungspotential ist **hoch** und die Umsetzung kann automatisiert erfolgen. Abgesehen von untertägigen Ereignisdauern kann die gesamte Information des Quellschemas in das Zielschema übertragen werden. Für Schadeinträge gibt es zum Teil separate Angaben zu Zeitpunkt und Dauer. Fehlen solche Angaben, wird im CESARE-Zielschema die im Original vorhandene Zeitinformation zum Ereignis auch für den dazugehörigen Schadeintrag übernommen.

#### HWFDB

In HWFDB-Datensatz beziehen sich fünf Merkmale auf den Ereigniszeitraum und die Zuverlässigkeit der Angaben zum Ereigniszeitraum:

- BEGINN: Datum und Uhrzeit des Ereignisbeginns
- DAUER\_TAGE: Tageskomponente der Angabe zur Ereignisdauer

- DAUER\_STUNDEN: Stundenkomponente der Angabe zur Ereignisdauer
- DAUER\_MINUTEN: Minutenkomponente der Angabe zur Ereignisdauer
- EREIGNISDAUERMAXO: der sich auf die Ereignisdauer beziehende MAXO-Wert

Das Harmonisierungspotential ist **hoch** und die Umsetzung kann automatisiert erfolgen. Abgesehen von der Ereignisuhrzeit und untertägigen Ereignisdauern kann die gesamte Information des Quellschemas in das Zielschema übertragen werden. Mithilfe der Angaben zur Dauer wird aus den Originalangaben das Ereignisende für das Zielschema berechnet. Für Schadeinträge werden die Datumsangaben des entsprechenden Ereignisses in das CESARE-Zielschema übernommen.

#### VIOLA

In VIOLA wird der Ereigniszeitraum sowie die Zuverlässigkeit der Angaben zum Ereigniszeitraum anhand von vier Merkmalen dokumentiert:

- beginn: Datum und Uhrzeit des Ereignisbeginns
- ende: Datum und Uhrzeit des Ereignisendes
- beginn\_delta: Unsicherheitsangabe zum Ereignisbeginn; sie reicht von +/- 1 Minute bis zu +/- 6 Monaten
- ende\_delta: Unsicherheitsangabe zum Ereignisende; sie reicht von +/- 1 Minute bis zu +/- 6 Monaten

Das Harmonisierungspotential ist **hoch** und die Umsetzung kann automatisiert erfolgen. Abgesehen von der Ereignisuhrzeit und dem genauen Ausmaß der Unsicherheit kann die gesamte Information des Quellschemas in das Zielschema übertragen werden. Die Ereignisdauer für das Zielschema wird aus den Originalmerkmalen „beginn“ und „ende“ abgeleitet. Die MAXO-Werte zur Ereignisdauer werden aufgrund der ex-post Ableitung auf „A“ (Annahme/Schätzung) gesetzt. Die Unsicherheitsangaben in den Original-Merkmalen „beginn\_delta“ und „ende\_delta“ fließen über die MAXO-Werte in das Zielschema ein. Hierzu wird der Beginnzeitpunkt um „beginn\_delta“ verändert. Unterscheidet sich zwischen resultierendem „Minimum“-Beginnzeitpunkt und „Maximum“-Beginnzeitpunkt das Jahr, werden im Zielschema die MAXO-Werte für Beginn\_Jahr, Beginn\_Monat und Beginn\_Tag jeweils auf „A“ statt gesetzt. Ist das Jahr ident und unterscheidet sich das Monat, wird der MAXO-Wert für Beginn\_Jahr auf „M“ gesetzt, während die MAXO-Werte für Beginn\_Monat und Beginn\_Tag ein „A“ erhalten. Sind Jahr und Monat ident, wird lediglich der MAXO-Wert für Beginn\_Tag auf „A“ gesetzt. Die MAXO-Codierung für den Endzeitpunkt erfolgt auf dieselbe Weise. Für Schadeinträge werden die Datumsangaben des entsprechenden Ereignisses in das CESARE-Zielschema übernommen.

#### Einsatzdaten Landesfeuerwehr NÖ

Im Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich ist der Einsatzzeitraum anhand von zwei Merkmalen festgehalten:

- Datum\_von: Datum und Uhrzeit des Einsatzbeginns
- Datum\_bis: Datum und Uhrzeit des Einsatzendes

Bei rund 30 % der Einträge im Datensatz fehlt das Einsatzende. Im Fall von fehlenden Werten wurde auf Empfehlung des Datenbereitstellers das Einsatzende auf Basis der durchschnittlichen Einsatzdauer je Meldebild nachträglich geschätzt. Um die mit der Schätzung einhergehende Unsicherheit im Zielschema abzubilden, wird

der MAXO-Wert zum Merkmal „Ende\_Tag“ auf „A“ (Annahme/Schätzung) gesetzt. Unterscheiden sich Jahr und/oder Monat des geschätzten Enddatums vom Einsatzbeginn, werden zusätzlich die MAXO-Werte für „Ende\_Jahr“ und/oder „Ende\_Monat“ ebenfalls auf „A“ – statt auf „M“ (Messung/Feststellung) – gesetzt.

Insgesamt ist das Harmonisierungspotential **hoch** und die Umsetzung kann automatisiert erfolgen. Abgesehen von der Einsatzzeit kann die gesamte Information des Quellschemas in das Zielschema übertragen werden. Die Einsatzdauer wird aus den Originalmerkmalen „Datum\_von“ und „Datum\_bis“ abgeleitet. Der MAXO-Wert zur Einsatzdauer richtet sich nach dem für das Enddatum vergebenen MAXO-Wert.

Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Wie im CESARE-Deliverable D4.1 beschrieben, liegt uns der Datensatz „KatFonds NÖ Privat“ derzeit nur in Jahresauflösung vor, d. h. für jeden Schadeintrag ist lediglich das Jahr des Schadeintritts bekannt. Das Harmonisierungspotential ist **hoch** und die Umsetzung kann automatisiert erfolgen. Die gesamte Information des Quellschemas kann dabei in das Zielschema übertragen werden. Im Zielschema können dadurch die Merkmale Beginn\_Jahr und Ende\_Jahr sowie die dazugehörigen MAXO-Merkmale (jeweils mit „M“) befüllt werden. Die übrigen Datumsmerkmale des Zielschemas bleiben leer.

Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Im Datensatz „KatFonds NÖ Gemeinden“ ist der Schadenzeitraum anhand von zwei Merkmalen festgehalten:

- Datum vom: Datum, das den Beginn des Schadeintritts markiert
- Datum bis: Datum, das das Ende des Schadeintritts markiert

Es gibt kein Original-Merkmal, das die Zuverlässigkeit des angeführten Schadenzeitraums beschreibt, allerdings scheint es z. T. eine gewisse Unschärfe in der Datumsangabe zu geben, die sich vor allem in langen Zeitspannen zwischen „Datum vom“ und „Datum bis“ ausdrücken kann. In manchen Fällen legt das Merkmal „Kurzbeschreibung“ nahe, dass es sich tatsächlich um längere Prozesse handelt (z. B. bei Eintrag 710: „über mehrere Monate Setzungen bei 2 Hauseinfahrten“). In anderen Fällen scheinen hingegen mehrere Ereignisse in einem Eintrag zusammengefasst zu sein (z. B. bei Eintrag 1155, wo sich der angeführte Schadenzeitraum von 27. 1. 2008 bis 1. 3. 2008 erstreckt und die Kurzbeschreibung wie folgt lautet: „Schäden durch Orkan ‚Paula‘ am 27. 1. 2008 und durch Orkan ‚Emma‘ am 1. 3. 2008“). Einträge, deren Zeitspanne und Kurzbeschreibung multiple Schadereignisse nahelegen, umfassen die fortlaufenden Zahlen 1109, 1155, 2244 und 2246. Sie werden im Zuge des Harmonisierungsprozesses getrennt, wobei eine gleichmäßig proportionale Aufteilung der Schäden erfolgt.

Bei sechs Einträgen des zur Verfügung gestellten Datensatzes liegt das Beginndatum hinter dem Enddatum. Hier scheint es sich um eine Vertauschung der Einträge zu handeln, die im Zuge des Harmonisierungsprozesses bereinigt wird. Dies betrifft folgende Einträge (Original-Merkmal „fortlaufende Zahl“): 19, 649, 1078, 1277, 2824 und 4464.

Insgesamt ist das Harmonisierungspotential **hoch** und die Umsetzung kann automatisiert erfolgen. Die gesamte Information des Quellschemas kann dabei in das Zielschema übertragen werden. Nachdem Angaben zur Zuverlässigkeit des Datums fehlen, setzen wir im Zielschema die MAXO-Werte für Jahres- und Monatsangaben grundsätzlich auf „M“ (Messung/Feststellung) und lassen jene für Tagesangaben leer. Beträgt die Spanne zwischen

Beginn und Ende mehr als fünf Tage, wird der MAXO-Wert für Ende\_Monat bzw. Ende\_Jahr ebenfalls leer gelassen, falls sich das Monat bzw. Jahr im Enddatum vom Monat bzw. Jahr im Beginndatum unterscheidet.

Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Im Datensatz „KatFonds Stmk Privat“ bezieht sich ein Merkmal auf das (taggenaue) Schadendatum. Das Harmonisierungspotential ist **hoch** und die Umsetzung kann automatisiert erfolgen. Das im Original angeführte Datum wird im Zielschema als Beginn- und Enddatum verwendet. Es gibt kein Original-Merkmal, das die Zuverlässigkeit des angeführten Schadenzeitraums beschreibt. Wir setzen daher im Zielschema die MAXO-Werte für Jahres- und Monatsangaben auf „M“ (Messung/Feststellung) und lassen jene für Tagesangaben leer.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Im Datensatz „KatFonds Stmk Land“ bezieht sich eines der Original-Merkmale auf das Schadendatum. Das Schadendatum ist darin in unterschiedlichsten Formaten und in unterschiedlichem Detailgrad (von jahresgenau bis taggenau) angeführt, z. B. „05.10.2005“, „Sept.-Okt. 2005“, „Sommer 2005“, „21.-22.08.2005“, „2005“. Zum Teil sind mehrere Schadereignisse in einem Schadeintrag zusammengefasst, z. B. „11.02. u. 27.04.2009“. In diesen Fällen erfolgt im Zuge des Harmonisierungsprozesses eine Aufteilung des Schadeintrags in die einzelnen Schadereignisse mit gleichmäßig proportionale Aufteilung der Schäden.

Insgesamt ist das Harmonisierungspotential **hoch**, die Harmonisierung kann allerdings aufgrund diverser Sonderfälle nur semi-automatisiert erfolgen. Jahreszeitangaben wie „Sommer“, „Winter“, „Herbst“ etc. werden meteorologisch interpretiert, wobei die mit der Bezeichnung einhergehende Unschärfe durch ein „A“ (Annahme/Schätzung) in den MAXO-Werten der Ziel-Merkmale Beginn\_Monat und Ende\_Monat dargestellt wird. Können fehlende Datumskomponenten im Zuge des Harmonisierungsprozesses im Nachhinein geschätzt werden, erfolgt eine entsprechende Kennzeichnung („A“) über den dazugehörigen MAXO-Wert. Dies ist in Einzelfällen beispielsweise bei Lawinenschäden möglich, wo nur Jahr und Monat bekannt sind und durch Abgleich mit dem WLK-Datensatz zum Teil auf den Tag des Schadeneintritts rückgeschlossen werden kann. Fehlende Datumskomponenten, die nachträglich nicht aufgefüllt werden können, werden hingegen mit einem „O“ (nicht bestimmbar) im entsprechenden MAXO-Merkmal gekennzeichnet. Bei vollständig vorhandenen Datumsangaben werden im Zielschema die MAXO-Werte für Jahres- und Monatsangaben auf „M“ (Messung/Feststellung) gesetzt, jene für Tagesangaben hingegen aufgrund der insgesamt im Originaldatensatz fehlenden Angaben zur Zuverlässigkeit des Datums leer gelassen.

**Verbleibende Herausforderungen im Bereich Datums-/Zeitangaben**

Durch vorhandene Unschärfen in den Datums- und Zeitangaben können unterschiedliche Datenquellen für ein und dasselbe Ereignis abweichende Datums- und Zeitangaben aufweisen. Diese Unschärfen lassen sich auch im Zuge des Harmonisierungsprozesses nicht beseitigen. Somit besteht durch die Zusammenführung unterschiedlicher Datensätze ein gewisses Risiko für Mehrfachzählungen von Ereignissen, Prozessen und Schäden. Wie in den folgenden Abschnitten noch genauer beschrieben (siehe v. a. Unterabschnitt „Schadausmaß“), werden insbesondere in Hinblick auf die Informationen zum Schadausmaß unterschiedliche Vorkehrungen getroffen, um das Risiko für Mehrfachzählungen möglichst gering zu halten.

## VERORTUNG

### Definition des CESARE Zielschemas

Wie im CESARE-Deliverable D4.1 beschrieben, variiert der Detailgrad der räumlichen Verortung in den zur Verfügung stehenden Originaldaten von Punkt- bzw. Linienverortungen (WLK, GEORIOS, HWFDB, Einsatzdaten Landesfeuerwehr) bis hin zu Polygonen auf Gemeinde-, Bezirks-, Regions- und Bundeslandebene (VIOLA, KatFonds-Datensätze). Im Sinne eines „kleinsten gemeinsamen Nenners“ wird in CESARE vorerst die Gemeindeebene als gemeinsame Zielvariable definiert. Bezirke und Bundesländer bestehen aus einzelnen Gemeinden und können daher ebenfalls als eine Menge an mehreren Gemeinden dargestellt werden. Im Zuge der Harmonisierung werden somit alle originalen Verortungen auf die Gemeindeebene zum Stand 2020 überführt. Detailliertere Originalverortungen werden aber weiterhin mitgeführt und auch im CESARE Demonstrator (siehe AP 5) zum Teil dargestellt. Die harmonisierten Merkmale in Bezug auf die Verortung umfassen somit:

- Gemeindename (enumeration)
- Gemeindenummer (enumeration)

Die harmonisierte Verortung auf Gemeindeebene (Stand 2020) wird als separate Tabelle geführt, die über einen eindeutigen Identifier mit den Ereignis-, Phänomen- oder Schadeinträgen verbunden ist. Auf diese Weise können beispielsweise Schadeinträge, die mehrere Gemeinden umfassen, vom Quell- ins Zielschema überführt werden, ohne vorerst eine Aufteilung der Schäden auf die betroffenen Gemeinden vornehmen zu müssen. Eine approximative Aufteilung (z. B. auf Basis der Gemeindegrößen oder -anzahl) ist damit erst im Falle von Auswertungen oder Darstellungen auf Gemeindeebene nötig.

### Harmonisierungspotential und Umsetzung je Datensatz

#### GEORIOS

Die in GEORIOS dokumentierten Ereignisse sind punktverortet, wobei ein- und dasselbe Ereignis zum Teil durch mehr als einen Punkt definiert ist (z. B. Verortung des Abriss- und des Akkumulationsbereichs). Die Harmonisierung von Punkt- auf Gemeindeebene bringt einen deutlichen Informationsverlust mit sich, nachdem die Zielvariable viel gröber aufgelöst ist als das Originalmerkmal. Aus Sicht der Zielvariable ist das Harmonisierungspotential im Sinne der Zuordnung der Punkte zu den zugrundeliegenden Gemeinden dennoch als **hoch** einzustufen. Die Harmonisierung kann automatisiert erfolgen.

#### WLK

Im WLK-Datensatz sind die dokumentierten Ereignisse ebenfalls punktverortet. Auch eine Zuweisung zu Gemeinde, Bezirk und Bundesland ist bereits im Original vorhanden, wird im Zuge des Harmonisierungsprozesses aber wiederholt, um die Gemeindestruktur zum Stand 2020 abzubilden. In Grenzregionen kann es dabei ebenfalls zu Neuordnungen kommen. Analog zum GEORIOS-Datensatz ist das Harmonisierungspotential aus Sicht der Zielvariable **hoch** und die Harmonisierung kann automatisiert erfolgen.

### HWFDB

Im HWFDB-Datensatz sind die dokumentierten Ereignisse punkt- und/oder linienverortet. Analog zum GEORIOS- und WLK-Datensatz ist das Harmonisierungspotential aus Sicht der Zielvariable **hoch** und die Harmonisierung kann automatisiert erfolgen. Linienverortungen werden dabei all jenen Gemeinden zugeordnet, durch die sie führen.

### VIOLA

In VIOLA sind die dokumentierten Ereignisse als Polygone verortet. Bei den Polygonen handelt es sich meist um ein- oder mehrere Gemeinden und Bezirke, aber zum Teil auch um Örtlichkeiten, Regionen oder ganze Bundesländer. Das Harmonisierungspotential ist aus Sicht der Zielvariable **hoch** und die Harmonisierung kann automatisiert erfolgen. Polygone, die nicht aus ein- oder mehreren Gemeinden bestehen, werden jenen Gemeinden zugeordnet, die mit ihnen eine Überschneidung aufweisen.

### Einsatzdaten Landesfeuerwehr NÖ

Die Einsatzdaten des Landesfeuerwehrverbandes Niederösterreich sind in der Regel punktverortet, jedoch mit unterschiedlicher Genauigkeit, die in einem separaten Merkmal dokumentiert ist. Liegt keine Punktverortung vor, sind weitere Ortsinformationen angeführt, die eine Gemeindezuordnung ermöglichen. Das Harmonisierungspotential ist damit aus Sicht der Zielvariable **hoch** und die Harmonisierung kann automatisiert erfolgen.

### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Im Datensatz „KatFonds NÖ Privat“ liegen die Schadeinträge bereits auf Gemeindeebene vor. Hier erfolgt im Rahmen des Harmonisierungsprozesses lediglich eine Aktualisierung auf den Stand 2020, wobei es im Falle von Niederösterreich zwischen 2005 und 2020 bis auf einige wenige Namensumstellungen keine Änderungen in der Gemeindestruktur gab (vgl. Statistik Austria, 2020). Das Harmonisierungspotential ist **hoch** und die Harmonisierung kann automatisiert erfolgen.

### Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Analog zu „KatFonds NÖ Privat“.

### Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Im Datensatz „KatFonds Stmk Privat“ liegen die Schadeinträge ebenfalls auf Gemeindeebene vor. Im Zuge des Harmonisierungsprozesses erfolgt damit lediglich eine Aktualisierung auf den Stand 2020. Anders als in Niederösterreich gab es in der Steiermark zwischen 2005 und 2020 deutliche Umstellungen in der Gemeindestruktur, v. a. im Zuge der Gemeindestrukturreform 2015 (vgl. Statistik Austria, 2020). Dabei kam es überwiegend zu Gemeindezusammenlegungen. Fünf Gemeinden wurden im Zuge dieser Zusammenlegungen allerdings getrennt und jeweils unterschiedlichen Neu-Gemeinden zugeordnet. Die Harmonisierung von der alten auf die neue Gemeindestruktur bringt einen gewissen Informationsverlust mit sich, nachdem die neue Struktur eine gröbere Auflösung aufweist als die alte. Wie auch im Falle der Punktverortungen ist das Harmonisierungspotential aus Sicht der Zielvariable aber als **hoch** einzustufen. Die Harmonisierung kann automatisiert erfolgen.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Auch im Datensatz „KatFonds Stmk Land“ liegen die Schadeinträge i. d. R. auf Gemeindeebene vor. Da im Originaldatensatz nur die Gemeindennamen – zum Teil in unterschiedlichster Schreibweise – und nicht auch die Gemeindekennzahlen geführt sind, erfolgt im Vorfeld der Harmonisierung eine Zuweisung der Gemeindekennzahlen zu den Gemeindennamen gemäß Stand zum jeweiligen Schadzeitpunkt. In den vereinzelt Fällen, in denen der Gemeindename unleserlich<sup>2</sup> ist oder in denen statt konkreter Gemeindennamen „mehrere Gemeinden“ bzw. „diverse Gemeinden“ angeführt ist, wird versucht, die betroffenen Gemeinden nachträglich über im Originaldatensatz mitgeführte Zusatzinformationen abzuleiten. Je nach geschädigtem Objekt kann es sich bei diesen Zusatzinformationen beispielsweise um den Namen des betroffenen Gewässers oder der betroffenen Landesstraße (z. T. mit Kilometerangabe) handeln. Zuordnungen von betroffenen Gewässern zu Gemeinden erfolgen auf Basis des Datensatzes „Fließgewässer des Gesamtwässernetz Österreich (GGN), Version v16“<sup>3</sup>, Zuordnungen von Landesstraßen auf Basis des Datensatzes „Übergeordnetes Straßennetz Steiermark (Stand 2015)“<sup>4</sup>. Der Harmonisierungsprozess selbst läuft analog zum Datensatz „KatFonds Stmk Privat“ ab.

**Verbleibende Herausforderungen im Bereich Verortung**

Die unterschiedlichen räumlichen Auflösungen, aber auch naturgemäße Unsicherheiten in vorhandenen Punktverortungen bergen für das Zusammenführen von Daten aus heterogenen Quellen das Risiko von Mehrfachzählungen. So lässt sich aufgrund von Unsicherheiten in der Verortung meist nicht eindeutig sagen, ob es sich bei einem zeitgleich und in räumlicher Nähe dokumentierten Murenabgang im GEORIOS- und im WLK-Datensatz um zwei separate Ereignisse oder um ein- und dasselbe Ereignis handelt. Im Falle der KatFonds-Datensätze, die für CESARE in ihrer kleinsten räumlichen Auflösung auf Gemeindeebene vorliegen, kann ebenfalls oft nicht eindeutig abgeklärt werden, ob beispielsweise bestimmte Vermurungsschäden in einer Gemeinde auf eine im GEORIOS- oder WLK-Datensatz dokumentierte Mure zurückzuführen sind oder durch eine (oder mehrere) andere Mure(n) verursacht wurden. Das Risiko von Mehrfachzählungen kann durch die Verortungsharmonisierung nicht reduziert oder beseitigt werden. Wie in den folgenden Abschnitten noch genauer beschrieben (siehe v. a. Unterabschnitt „Schadausmaß“), werden insbesondere in Hinblick auf die Informationen zum Schadausmaß unterschiedliche Vorkehrungen getroffen, um das Risiko für Mehrfachzählungen möglichst gering zu halten.

**ART DER NATURGEFAHR (HAZARD)**

**Definition des CESARE Zielschemas**

Was die Art der Naturgefahr betrifft, unterscheiden sich die zur Verfügung stehenden Datenquellen in zweierlei Hinsicht: (1) in Bezug auf das verwendete Vokabular und (2) in Bezug auf die Art und Weise, wie die Klassifikation

---

<sup>2</sup> Die Originaldaten liegen zum überwiegenden Teil als Scans im pdf-Format vor und wurden im Rahmen von CESARE digitalisiert. Die Digitalisierung erfolgte – wo möglich – automatisiert, vielfach musste aufgrund der Bildqualität allerdings ergänzend oder vollständig auf manuelle Übertragung zurückgegriffen werden.

<sup>3</sup> <https://www.data.gv.at/katalog/dataset/gesamtwassernetzflussgewasserrouten/resource/fd9b49b2-039b-468f-bec5-2c1a53f9250e>

<sup>4</sup> [https://www.data.gv.at/katalog/dataset/land-stmk\\_bergeordnetesstraennetz](https://www.data.gv.at/katalog/dataset/land-stmk_bergeordnetesstraennetz)

an sich erfolgt, nämlich über den Prozess oder über das Ereignis, das dem Prozess zugrunde liegt bzw. ihn ausgelöst hat. So kann etwa ein durch ein Hochwasser ausgelöster Erdbeben in einem Datensatz als Erdbeben aufscheinen, im anderen hingegen als Hochwasser.

Das standardisierte Vokabular für das CESARE-Zielschema, das drei Ebenen unterschiedlicher Detailliertheit umfasst, ist ausführlich im CESARE-Deliverable D4.1 beschrieben und wird daher an dieser Stelle nicht näher erläutert. Für die Klassifikation an sich sieht das CESARE-Zielschema eine Zuordnung über den Prozess und nicht über das dem Prozess zugrundeliegende bzw. das den Prozess auslösende Ereignis vor. Zusätzlich zur Naturgefahrenart mit ihren drei Ebenen an Detailliertheit umfasst das CESARE-Zielschema auch ein MAXO-Merkmal zur Naturgefahrenart:

- Hazard\_1 (enumeration)
- Hazard\_2 (enumeration)
- Hazard\_3 (enumeration)
- Hazard\_MAXO (enumeration: {M, A, X, O})

Der MAXO-Wert zum Hazard wird im Zuge der Harmonisierung nur befüllt, wenn im Bearbeitungsprozess – abgesehen vom Mapping zwischen Quell- und Zielschema – zusätzliche Änderungen an der im Original genannten Naturgefahr vorgenommen werden oder es Gründe gibt, den im Original angeführten Wert als „unsicher“ zu erachten. In all diesen Fällen erfolgt die Kennzeichnung der einhergehenden Unsicherheit mit einem „A“ (Annahme/Schätzung).

#### Harmonisierungspotential und Umsetzung je Datensatz

##### GEORIOS

Im GEORIOS-Datensatz erfolgt die Klassifikation der dort dokumentierten Massenbewegungsereignisse, so wie im CESARE-Zielschema vorgesehen, über den Prozess. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Da auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfachzuordnungen möglich bzw. nötig sind, wird das Harmonisierungspotential insgesamt als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert wird nicht befüllt.

##### WLK

Im WLK-Datensatz erfolgt die Klassifikation der dokumentierten Ereignisse, so wie im CESARE-Zielschema vorgesehen, über den Prozess. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Da auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfachzuordnungen möglich bzw. nötig sind, wird das Harmonisierungspotential insgesamt als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert wird nicht befüllt.

### HWFDB

Der HWFDB-Datensatz enthält ausschließlich Hochwasserereignisse. Im zur Verfügung gestellten Auszug sind darüber hinaus keine weiteren Informationen zu bestimmten Hochwasserkategorien oder Prozessarten enthalten, weshalb alle im Datensatz gelisteten Ereignisse auf das Zielvokabular „Hochwasser“ (Ebene 2) bzw. „Hochwasser (unspezifiziert)“ (Ebene 3) gemappt werden (siehe auch CESARE-Deliverable D4.1). Das Harmonisierungspotential wird als **hoch** eingestuft und die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert wird nicht befüllt.

### VIOLA

Im VIOLA-Datensatz erfolgt die Klassifikation der dokumentierten Ereignisse, so wie im CESARE-Zielschema vorgesehen, über den Prozess. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Insgesamt kann das Harmonisierungspotential als **hoch** eingestuft werden und die Harmonisierung automatisiert erfolgen. Der MAXO-Wert wird nicht befüllt.

### Einsatzdaten Landesfeuerwehr NÖ

Im Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich lässt das Original-Merkmal „Meldebild“ Aufschlüsse über die einsatzverursachende Art der Naturgefahr zu. Allerdings ist das „Meldebild“ insgesamt mit Unsicherheiten behaftet, da es nach Auskunft des Datenbereitstellers in der Vergangenheit keine klare Abgrenzung zwischen den Meldebildern gab. Vielmehr hing es von den handelnden Disponenten ab, welches Meldebild auf Basis der Beschreibung des Notruf-Melders bzw. der Notruf-Melderin ausgewählt wurde. Vor allem zwischen den Meldebildern „Auspumparbeiten“, „Unwettereinsatz“ und „Hochwasser“ dürfte die Abgrenzung fließend verlaufen. Aufgrund dieser Unsicherheiten wird das Harmonisierungspotential insgesamt als **mittel** eingestuft. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Die Harmonisierung an sich kann automatisiert erfolgen. Der MAXO-Wert wird aufgrund der vorliegenden Unsicherheiten durchgehend auf „A“ (Annahme/Schätzung) gesetzt.

### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Im Datensatz „KatFonds NÖ Privat“ erfolgt die Klassifikation der schadenverursachenden Naturgefahr, so wie im CESARE-Zielschema vorgesehen, über den Prozess. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Da auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfachzuordnungen möglich bzw. nötig sind, wird das Harmonisierungspotential insgesamt als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert wird nicht befüllt.

### Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Im Datensatz „KatFonds NÖ Gemeinden“ erfolgt die Klassifikation der schadenverursachenden Naturgefahr für gewöhnlich nicht über den Prozess, sondern über das Ereignis, das dem schadenverursachenden Prozess zugrunde liegt. Eine durch ein Hochwasser ausgelöste Vermurung wird demnach tendenziell nicht unter Vermurung, sondern unter Hochwasser geführt. Selbiges gilt etwa für Erdbeben. Insbesondere die Abgrenzung zwi-

schen Hochwasser und Massenbewegung unterscheidet sich damit vom Zielschema. Eine nachträgliche Klassifizierung nach Prozess anstatt nach zugrundeliegendem Ereignis anhand des Original-Merkmals „Kurzbeschreibung“ gestaltet sich als schwierig, da laut Kurzbeschreibung oftmals mehrere Prozesse in einem Eintrag zusammengefasst sind. Die Informationen aus der Kurzbeschreibung fließen aber, wie im Folgenden noch näher beschrieben, in die Festlegung des MAXO-Werts ein. Aufgrund der unterschiedlichen Klassifikationsweise im Quell- und Zielschema und dem Umstand, dass auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfachzuordnungen möglich bzw. nötig sind (siehe CESARE-Deliverable D4.1), wird das Harmonisierungspotential insgesamt als **mittel** eingestuft. Die Harmonisierung des Vokabulars kann automatisiert erfolgen. Wenn die im Original-Merkmal „Schadensursache“ dokumentierte schadenverursachende Naturgefahr von dem in der Kurzbeschreibung genannten Prozess abweicht oder in der Kurzbeschreibung noch weitere Prozesse genannt sind, wird der dazugehörige MAXO-Wert zum Ausdruck der Unsicherheit auf „A“ (Annahme/Schätzung) gesetzt. In zwei Fällen (fortlaufende Nummer: 1095, 1155) wurde die dokumentierte schadenverursachende Naturgefahr aufgrund einer widersprüchlichen Kurzbeschreibung im Zuge des Harmonisierungsprozesses abgeändert. Auch hier wird der dazugehörige MAXO-Wert auf „A“ gesetzt. Für die restlichen Einträge wird der MAXO-Wert nicht befüllt.

Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Im Datensatz „KatFonds Stmk Privat“ erfolgt die Klassifikation der schadenverursachenden Naturgefahr, so wie im CESARE-Zielschema vorgesehen, über den Prozess. Das Mapping zwischen Quell- und Zielvokabular ist im CESARE-Deliverable D4.1 dargestellt. Da auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfachzuordnungen möglich bzw. nötig sind, wird das Harmonisierungspotential insgesamt als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert wird grundsätzlich nicht befüllt. Eine Ausnahme bilden Einträge, die keiner schadenverursachenden Naturgefahr zugeordnet sind. Dies betrifft 1,3 % der Einträge des Datensatzes. Hier wird der MAXO-Wert auf „O“ (nicht bestimmbar) gesetzt. Eine weitere Ausnahme stellen Einträge dar, die als Original-Schadenart die Kategorie „Schäden an privaten Grundstücken und Gebäuden, die durch Erdbeben entstanden sind und durch Tiefendrainagen und Sicherungen an Gebäudefundamenten behoben werden müssen“ aufweisen, jedoch als Original-Schadenursache nicht wie zu erwarten die Kategorie „Erdbeben“, sondern eine andere Naturgefahrenart listen. In diesen Fällen wird der MAXO-Wert auf „A“ (Annahme/Schätzung) gesetzt. Ebenfalls auf „A“ gesetzt wird der MAXO-Wert für jene zwei Einträge, deren originale Schadenursache im Zuge des Harmonisierungsprozesses aufgrund von Ungeheimheiten in der Kombination Datum, Lokalisation und Schadensursache nachträglich abgeändert wurde.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Im Datensatz „KatFonds Stmk Land“ erfolgt die Klassifikation der schadenverursachenden Naturgefahr, so wie im CESARE-Zielschema vorgesehen, über den Prozess, allerdings kommen häufig Mehrfachnennungen vor (~ 15 % der Einträge). Pro Schadeintrag können bis zu drei schadenverursachende Naturgefahren genannt sein. Für die Harmonisierung wird jeweils nur die erstgenannte Naturgefahr herangezogen, jedoch wird die mit Mehrfachnennungen einhergehende Unsicherheit durch ein „A“ (Annahme/Schätzung) im zugehörigen MAXO-Merkmal widerspiegelt. Aufgrund der vorhandenen Unsicherheit durch Mehrfachnennungen und dem Umstand, dass auf der dritten Detailebene bei der Zuordnung des Quellvokabulars zum Zielvokabular zum Teil Mehrfach-

zuordnungen möglich bzw. nötig sind (siehe CESARE-Deliverable D4.1), wird das Harmonisierungspotential insgesamt als **mittel** eingestuft. Die Harmonisierung des Vokabulars kann automatisiert erfolgen. Der MAXO-Wert wird grundsätzlich nicht befüllt, außer beim Vorhandensein von Unsicherheiten z. B. durch Mehrfachnennungen („A“) und bei fehlender Schadenursache („O“).

### Verbleibende Herausforderungen im Bereich Naturgefahrenart

Während sich das Vokabular grundsätzlich gut harmonisieren lässt, stellen sowohl die zum Teil in den Originaldaten vorhandenen Unschärfen in Bezug auf die Abgrenzung zwischen Naturgefahren als auch die teilweise vom Zielschema abweichende Klassifikation der Ereignisart eine verbleibende Herausforderung dar. Unschärfen in der Abgrenzung sind beispielsweise im Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich zu finden, aber auch in den „KatFonds“-Datensätzen. Im Datensatz „KatFonds NÖ Gemeinde“ weicht zudem die Klassifikation der Ereignisart vom Zielschema ab, indem Schäden nicht dem verursachenden Prozess, sondern dem Ereignis zugeordnet werden, der dem schadenversuchenden Prozess zugrunde liegt. Auch im Zielschema bleiben die Informationen rund um die Naturgefahrenart somit zum Teil mit Unsicherheiten behaftet, was über das dazugehörige MAXO-Merkmal jeweils zum Ausdruck gebracht wird.

## SCHADENKATEGORIE BZW. ART DES GESCHÄDIGTEN ELEMENTS

### Definition des CESARE Zielschemas

Bei der Schadenkategorie bzw. Art des geschädigten Elements unterscheiden sich die zur Verfügung stehenden Datenquellen in Bezug auf das verwendete Vokabular. Die Harmonisierung erfolgt durch Zusammenführen in ein gemeinsames, standardisiertes Vokabular. Das standardisierte Vokabular für das CESARE-Zielschema zur Schadenkategorie bzw. Art des geschädigten Elements ist im CESARE-Deliverable D4.1 bereits ausführlich beschrieben. Es besteht aus drei Ebenen unterschiedlicher Detailliertheit, wobei auf der obersten Ebene zwischen Personenschaden, Sachschaden und Einsatzkosten unterschieden wird. Insgesamt umfasst das CESARE-Zielschema somit folgende Merkmale:

- Schadenkategorie\_0 (enumeration)
- Schadenkategorie\_1 (enumeration)
- Schadenkategorie\_2 (enumeration)

### Harmonisierungspotential und Umsetzung je Datensatz

#### GEORIOS

Wie im CESARE-Deliverable D4.1 beschrieben, sind etwaige Angaben zu Schäden im GEORIOS-Datensatz in Form eines Freitextes gespeichert. Vorerst werden für CESARE nur Personenschäden aus dem Freitext extrahiert. Das Mapping zwischen Quell- und Zielvokabular, dessen Anwendung semi-automatisiert erfolgt, ist im CESARE-Deliverable D4.1 dargestellt. Das Harmonisierungspotential ist **hoch**.

### WLK

Geschädigte Elemente werden im WLK-Datensatz auf vier unterschiedlichen Detailebenen kategorisiert. Auf der obersten Ebene werden Personenschäden, Sach-/Nutzungsschäden, Schäden an Tieren und Naturraumschäden unterschieden. Das Mapping zwischen Quell- und Zielvokabular, dessen Anwendung automatisiert erfolgt, ist im CESARE-Deliverable D4.1 dargestellt. Das Harmonisierungspotential wird aufgrund nicht ausschließbarer semantischer Unschärfen als **mittelhoch** eingestuft.

### HWFDB

Aus dem HWFDB-Datensatz gehen Schäden an Hochwasser-Schutzbauten in das CESARE-Zielschema zu geschädigten Elementen Schäden ein (siehe auch CESARE-Deliverable D4.1). Das Harmonisierungspotential ist **hoch**.

### VIOLA

Im VIOLA-Datensatz werden zwei Detailebenen in Bezug auf die Schadenkategorie geführt. Die oberste Ebene unterscheidet dabei zwischen Einsatzkräften, Sachschäden, Flurschäden und Personenschäden. Das Mapping zwischen Quell- und Zielvokabular, dessen Anwendung automatisiert erfolgt, ist im CESARE-Deliverable D4.1 dargestellt. Das Harmonisierungspotential wird aufgrund nicht ausschließbarer semantischer Unschärfen als **mittelhoch** eingestuft.

### Einsatzdaten Landesfeuerwehr NÖ

Im Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich bezieht sich jeder Eintrag auf einen Einsatz der Feuerwehr (siehe auch CESARE-Deliverable D4.1) und wird daher in die entsprechende Zielkategorie überführt. Das Harmonisierungspotential ist **hoch**.

### Dokumentationen der Landesverwaltungen (KatFonds-Datensätze)

Wie im CESARE-Deliverable D4.1 beschrieben, verfügen die meisten zur Verfügung gestellten Datensätze der Landesverwaltungen („*KatFonds Stmk Privat*“, „*KatFonds NÖ Privat*“, „*KatFonds NÖ Gemeinden*“) in Bezug auf das geschädigte Element über vorgegebene Kategorien. Im Datensatz „*KatFonds Stmk Privat*“ handelt es sich hingegen teilweise auch um Freitexte. Das Mapping zwischen Quell- und Zielvokabular, dessen Anwendung je nach Datensatz automatisiert oder semi-automatisiert erfolgt, ist im CESARE-Deliverable D4.1 dargestellt. Großteils lassen sich die Originalkategorien eindeutig in die Zielkategorien überführen, in Ausnahmefällen ist jedoch auch die Funktion des Splittens notwendig, was mit gewissen Unschärfen einhergeht. Das Harmonisierungspotential wird aufgrund der Unschärfen durch das Splitten und weiterer nicht ausschließbarer semantischer Unschärfen als **mittelhoch** eingestuft.

### **Verbleibende Herausforderungen im Bereich Schadenkategorie**

Es lässt sich nur schwer beurteilen, wie konsistent in den Originaldaten die Zuordnung geschädigter Elemente zu den jeweils verwendeten Schadenkategorien erfolgt bzw. wie scharf die Abgrenzungen zwischen den einzelnen Kategorien verlaufen. Im Datensatz „*KatFonds NÖ Gemeinden*“ legt ein Vergleich zwischen dem Merkmal „Schadensobjekt“ und dem Merkmal „Kurzbeschreibung“ nahe, dass es in manchen Fällen zu Inkonsistenzen

kommen kann bzw. Abgrenzungen zum Teil etwas verschwimmen. Werden in der Kurzbeschreibung „FF Häuser“ (Gebäude der Freiwilligen Feuerwehr) genannt, finden sich diese im Merkmal „Schadenobjekt“ zum Teil unter der Kategorie „Amtshäuser“ und zum Teil unter der Kategorie „allg. Gebäude/baul. Anlagen“ wieder. In der Kurzbeschreibung vorkommende „Teichanlagen“ sind im Merkmal „Schadenobjekt“ teilweise der Kategorie „Teichanlagen“ und teilweise der Kategorie „Sport- und Freizeitanlagen“ zugeordnet, je nachdem, wo ihr Nutzungsschwerpunkt liegt. Etwaige Inkonsistenzen und verschwimmende Abgrenzungen können auch für die anderen Datensätze nicht ausgeschlossen werden.

## SCHADAUSMAß

### Definition des CESARE Zielschemas

Wie in Deliverable D4.1 beschrieben, liegt der Fokus in CESARE vorerst auf Schadenindikatoren zu menschlichen Verlusten (Personenschäden) sowie zu direkten wirtschaftlichen Verlusten in Form von monetären Sachschäden und Einsatzkosten. Indikatoren zu Sachschäden in physischen Einheiten (z. B. Anzahl beschädigter/zerstörter Häuser, Kilometer beschädigter Straßen, Hektar beschädigter landwirtschaftlicher Flächen etc.) werden im Demonstrator vorerst nicht berücksichtigt, da aufgrund der unterschiedlichen Datenquellen keine Konsistenz zwischen physisch und monetär gemessenen Schäden innerhalb ein und derselben betroffenen Elementkategorie hergestellt bzw. gewährleistet werden kann. Das Zielschema zum Schadausmaß umfasst derzeit fünf Merkmale:

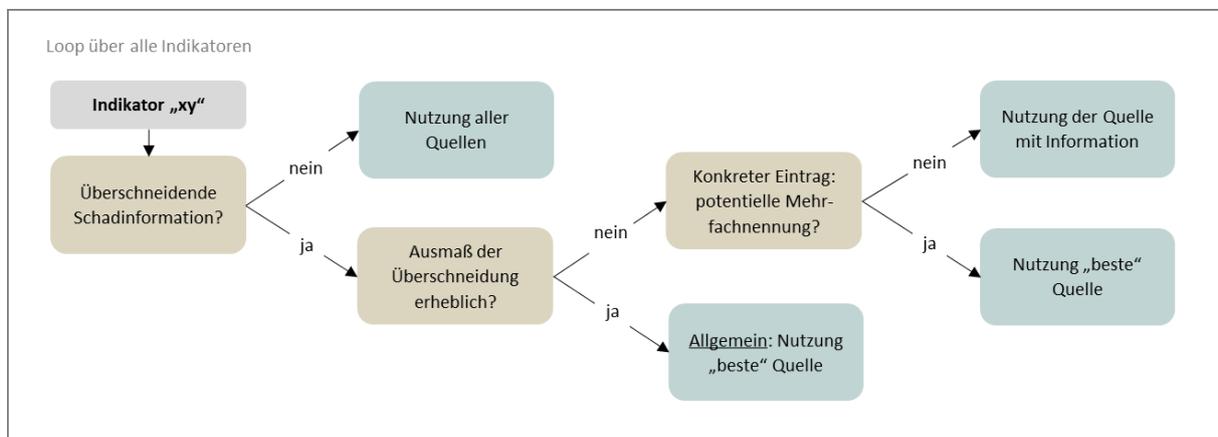
- Schadensumme (numeric)
- Schadensumme\_MAXO (enumeration: {M, A, X, O})
- Schaden\_Anzahl (integer)
- Schaden\_Anzahl\_MAXO (enumeration: {M, A, X, O})
- Doublerisiko (boolean)

Unter „Schadensumme“ wird das Ausmaß aller in monetären Einheiten gemessenen Schäden dokumentiert. Je nach Schadenkategorie handelt es sich um Wiederherstellungskosten für Sachschäden (= Zeitwert) oder um Einsatzkosten. Die Maßeinheit beträgt Euro zu laufenden Preisen. Basierend auf der Praxis der verfügbaren KatFonds-Datensätze, die, wie im Folgenden noch näher beschrieben, als Hauptquelle für Wiederherstellungskosten von Sachschäden herangezogen werden, verstehen sich die Wiederherstellungskosten im Zielschema als exklusive Umsatzsteuer, wenn es sich beim Besitzer bzw. der Besitzerin des geschädigten Objekts um eine vorsteuerabzugsberechtigte Einrichtung handelt, und andernfalls als inklusive Umsatzsteuer. Unter „Schaden\_Anzahl“ wird hingegen das Schadausmaß im Falle von Personenschäden angeführt. Das Merkmal ließe sich zum Teil auch zur Dokumentation von Sachschäden in physischen Einheiten verwenden (z. B. Anzahl beschädigter/zerstörter Häuser), wovon wie erwähnt vorerst aber abgesehen wird. Für beide Merkmale, Schadensumme und Schadenanzahl, wird der Grad der Zuverlässigkeit der einzelnen Einträge mittels MAXO-Code angegeben.

Das Zielschema zum Schadausmaß sieht keinen Mindestschaden für die Berücksichtigung eines Schadens im Demonstrator vor und schließt auch keine Schadensarten aus, solange es sich um direkte Schäden handelt. Indirekte Schäden, wie beispielsweise Umsatzeinbußen, werden hingegen (vorerst) nicht berücksichtigt.

Unschärfen bei den Datums- und Zeitangaben sowie der räumlichen Verortung, unterschiedliche räumliche Auflösungen und Unterschiede in der Klassifikation der schadenverursachenden Naturgefahr erschweren bei der Zusammenführung unterschiedlicher Datenquellen die Identifikation identer Ereignisse und erhöhen damit auch das Risiko von Mehrfachzählungen im Bereich des Schadausmaßes. Um das Risiko von Mehrfachzählungen beim Schadausmaß dennoch möglichst gering zu halten, findet schon im Vorfeld ein Selektionsprozess statt, welche Informationen aus welchem Datensatz berücksichtigt werden (siehe auch Abbildung 3-1). Wo immer sich Schadinformationen aus unterschiedlichen Datenquellen nicht komplementär ergänzen, sondern potentiell überschneiden, wird im Vorfeld bewertet, welcher Datensatz je Schadenindikator die umfassenderen und/oder vertrauenswürdigeren Informationen liefert. Ist für den jeweils betrachteten Schadenindikator das Ausmaß der potentiellen Überschneidung durch die Zusammenführung heterogener Quellen erheblich, wird für diesen Schadenindikator von Haus aus nur jene Datenquelle herangezogen, die die umfassendsten und/oder vertrauenswürdigsten Informationen liefert. Ist das Ausmaß der potentiellen Überschneidung für den betrachteten Schadenindikator hingegen gering (im Sinne von nur wenigen Fällen, in denen dasselbe Ereignis bzw. derselbe Schaden durch mehrere Quellen erfasst ist), kann der Schadenindikator grundsätzlich durch mehrere Quellen bedient werden. In diesem Fall wird für jene Einträge, in denen Zeit, Raum und Naturgefahr eine akute potentielle Überschneidung nahelegen, jedoch nur die Schadinformation jener Quelle herangezogen, die als die umfassendste und/oder vertrauenswürdigste anzusehen ist. Die entsprechenden Einträge bzw. Schadinformationen in den restlichen für den betrachteten Schadenindikator herangezogenen Datenquellen werden über das CESARE-Zielmerkmal „Doublerisiko“ hingegen als potentielle Mehrfachzählungen markiert, sodass sie für Auswertungen ausgeschlossen werden können.

Abbildung 3-1: Selektionsprozess bei potentiell überschneidenden Schadinformationen aus heterogenen Quellen in Form eines Entscheidungsbaums



### Indikator „Personenschaden“

Informationen zu Personenschäden sind in den Datensätzen GEORIOS, WLK und VIOLA enthalten. Das Risiko von Mehrfachnennungen durch sich überschneidende Schadinformationen ist grundsätzlich gegeben, aber das Ausmaß der Überschneidung ist eher gering. Die Anzahl der Fälle, wo die Überschneidung von Zeitpunkt, Lokation und Naturgefahr auf das möglicherweise selbe Ereignis hinweisen und zusätzlich in mehr als einem der Datensätze ein dokumentierter Personenschaden vorliegt, sind überschaubar. Daher werden grundsätzlich alle drei Quellen zur Bedienung des CESARE-Indikators „Personenschaden“ herangezogen. Bei Schadeinträgen, deren

Zeitpunkt, Lokation und Ursache (Naturgefahr) auf eine potentielle Mehrfachnennung hinweisen, wird hingegen nur auf die umfassendste und/oder verlässlichste Quelle zurückgegriffen, während die entsprechenden Schadeinträge in den anderen Quellen anhand des Merkmals „Doublersisiko“ als potentielle Mehrfachnennungen markiert werden. Dazu wird überprüft, ob es in mindestens zwei der Datenquellen zum selben Startdatum (+/- 1 Tag) in derselben Gemeinde einen Personenschaden derselben Unterkategorie (Tote vs. Verletzte vs. Evakuierte) und verursacht durch dieselbe bzw. eine „artverwandte“ Naturgefahr<sup>5</sup> gab.

Folgende Priorisierung der Datenquellen kommt im Falle von Überschneidungen zur Anwendung:

- **Priorität 1: VIOLA**  
VIOLA basiert auf Informationen aus Medienberichten. Zwar wird diesen Informationen grundsätzlich keine höhere Zuverlässigkeit als den Informationen im WLK-Datensatz und in GEORIOS eingeräumt, jedoch beziehen sich die Einträge in VIOLA meist auf größere räumliche Einheiten und können daher theoretisch weitere Personenschäden außerhalb der sich konkret überschneidenden Gemeinde einschließen, die in den anderen Datensätzen nicht erfasst sind. Im Falle potentieller Überschneidungen wird im Falle von Personenschäden daher prioritär auf das in VIOLA angeführte Schadausmaß zurückgegriffen.
- **Priorität 2: WLK**  
Da in GEORIOS häufig Angaben zur konkreten Anzahl der geschädigten Personen fehlen, werden im Falle potentieller Überschneidungen Angaben zum Ausmaß von Personenschäden aus dem WLK-Datensatz gegenüber dem GEORIOS-Datensatz priorisiert.
- **Priorität 3: GEORIOS**  
Das Ausmaß von Personenschäden gemäß GEORIOS-Datensatz wird im Falle von Auswertungen nur dann berücksichtigt, wenn keine potentiellen Mehrfachnennungen mit dem WLK- oder VIOLA-Datensatz vorliegen.

#### Indikator „Sachschaden“

Informationen zum monetären Ausmaß von Sachschäden finden sich vor allem in den Datensätzen der Landesverwaltungen zu Katastrophenschäden (KatFonds-Datensätze), aber auch im WLK-, HWFDB- und VIOLA-Datensatz. Zum Teil sind die Informationen komplementär, häufig jedoch potentiell überschneidend. Dort, wo die Informationen potentiell überschneidend sind, ist das Ausmaß der potentiellen Überschneidung als erheblich einzustufen. In diesen Fällen wird von Haus aus nur jene Datenquelle herangezogen, die die umfassendsten und/oder vertrauenswürdigsten Informationen liefert.

Die Dokumentationen der Landesverwaltungen umfassen beihilfenfähige Katastrophenschäden zu Wiederherstellungskosten, die von Sachverständigen und Schadenerhebungskommissionen bewertet werden oder mittels Rechnungen zu belegen sind (siehe auch CESARE-Deliverable D4.1). Nicht enthalten, da nicht beihilfenfähig, sind

---

<sup>5</sup> Da es in den unterschiedlichen Datenquellen zum Teil zu unterschiedlichen Zuordnungen in Bezug auf die Naturgefahr kommen kann, wird dieser potentiellen Unschärfe Rechnung getragen, indem bei der Identifizierung potentieller Mehrfachnennungen zwischen ähnlichen Naturgefahren bzw. Naturgefahren, deren Abgrenzung je nach Datensatz fließend verlaufen könnte, nicht unterschieden wird.

u. a. versicherte (bzw. versicherbare) Schäden. Gegenüber dem VIOLA-Datensatz, der auf Medienberichten beruht und daher i. d. R grobe, vorläufige Schadensschätzungen enthält, sind die Informationen der KatFonds-Datensätze als räumlich detaillierter und robuster einzustufen, wenn auch u. U. zum Teil weniger umfassend (z. B. im Falle erforderlicher Mindestschäden für Dokumentation). Insgesamt wird daher in Bezug auf monetäre Sachschäden den KatFonds-Datensätzen gegenüber dem VIOLA-Datensatz der Vorrang gegeben.

Der WLK-Datensatz und die HWFDB enthalten neben Schätzungen zu den Wiederherstellungskosten von Schäden an Schutzbauten und Verbauungen der Wildbach- und Lawinerverbauung (Bundesebene) bzw. der Bundeswasserbauverwaltung zum Teil auch grobe Schätzungen zu Gesamtschäden (inkl. Privateigentum, Industrie, Infrastruktur etc.). Auch hier ist davon auszugehen, dass die KatFonds-Datensätze die robusteren, wenngleich u. U. zum Teil weniger umfassenden Zahlen liefern. Informationen zu den Wiederherstellungskosten von Schäden an Schutzbauten und Verbauungen der Wildbach- und Lawinerverbauung (Bundesebene) bzw. der Bundeswasserbauverwaltung werden hingegen aus dem WLK- und dem HWFDB-Datensatz übernommen. Um hier potentielle Überschneidungen mit den KatFonds-Datensätzen zu vermeiden, werden aus dem Datensatz „KatFonds Stmk Land“ bei Schäden im Bereich der Wildbach- und Lawinerverbauung oder des Schutzwasserbaus nur etwaige Landes- und Gemeindeanteile, aber keine Bundesanteile in das CESARE-Zielschema übernommen.

#### Indikator „Einsatzkosten“

Informationen zum monetären Ausmaß von Einsatzkosten finden sich im Datensatz „KatFonds Stmk Land“ in Form von Kosten für Soforthilfemaßnahmen wie beispielsweise Sicherungsmaßnahmen bzw. provisorische Maßnahmen zur Verhinderung einer Gefährdung von Leib und Leben und/oder bedeutenden Sachwerten (siehe auch CESARE-Deliverable D4.1). Mit den derzeit zur Verfügung stehenden anderen Datenquellen gibt es keine Überschneidungen. Der Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich und der VIOLA-Datensatz enthalten zwar Informationen darüber, wann und wo es zu Feuerwehreinsätzen bzw. Einsätzen von Notfallorganisationen gekommen ist, jedoch ist kein monetäres Schadausmaß in Form von Einsatzkosten aus den beiden Datensätzen ableitbar.

#### Harmonisierungspotential und Umsetzung je Datensatz

##### GEORIOS

Im GEORIOS-Datensatz finden sich vereinzelt Angaben zum Ausmaß von Personenschäden (extrahierbar aus dem Freitext zu Schäden). Das Harmonisierungspotential ist grundsätzlich **hoch** und die Harmonisierung kann automatisiert erfolgen. Ist nur bekannt, dass Personen geschädigt wurden aber nicht, wie viele, wird der dazugehörige MAXO-Wert auf „O“ (nicht bestimmbar) gesetzt. Ist die konkrete Anzahl der geschädigten Personen bekannt, wird der dazugehörige MAXO-Wert auf „M“ (Messung/Feststellung) gesetzt. Bei Einträgen, die potentielle Überschneidungen mit einer der anderen Datenquellen zu Personenschäden aufweisen (WLK, VIOLA), wird das in GEORIOS angeführte Schadausmaß zur Vermeidung von Mehrfachzählungen in Auswertungen nicht berücksichtigt.

##### WLK

Der WLK-Datensatz enthält (i) Angaben zum Ausmaß von Personenschäden, (ii) Angaben zum Ausmaß von Sachschäden in physischen Einheiten, (iii) grobe Schätzungen zum monetären Gesamtschaden (inkl. Privateigentum,

Industrie, Infrastruktur etc.), (iv) Schätzungen zu Wiederherstellungskosten von Schäden an WLV-Verbauungen (Bundesebene).

- Personenschäden: Angaben zum Ausmaß von Personenschäden werden in das CESARE-Zielschema übernommen. Das Harmonisierungspotential ist grundsätzlich **hoch** und die Harmonisierung kann automatisiert erfolgen. MAXO-Werte zum Schadausmaß sind bereits im Original vorhanden und werden in das CESARE-Zielschema übernommen. Bei Einträgen, die potentielle Überschneidungen mit Personenschäden aufweisen, die im VIOLA-Datensatz dokumentiert sind, wird das im WLK-Datensatz angeführte Schadausmaß zur Vermeidung von Mehrfachzählungen in Auswertungen nicht berücksichtigt.
- Sachschäden in physischen Einheiten: Informationen zum Ausmaß von Sachschäden in physischen Einheiten werden derzeit, wie erwähnt, nicht vom CESARE-Zielschema erfasst.
- Sachschäden (gesamt) in monetären Einheiten: Die (groben) Schätzungen zum monetären Gesamtschaden werden nicht in das CESARE-Zielschema übernommen, da ein hohes Risiko von Überschneidung und daraus resultierenden Mehrfachnennungen mit den KatFonds-Datensätzen besteht.
- Schäden an WLV-Verbauung: Schätzungen zu Wiederherstellungskosten von Schäden an WLV-Verbauungen (Bundesebene) werden in das CESARE-Zielschema zum Schadausmaß übernommen. Das Harmonisierungspotential ist **hoch** und die Harmonisierung kann automatisiert erfolgen. Der dazugehörige MAXO-Wert wird jeweils auf „A“ (Annahme/Schätzung) gesetzt.

#### HWFDB

Der HWFDB-Datensatz enthält (i) grobe monetäre Schätzungen zum Gesamtschaden (inkl. Privateigentum, Industrie, Infrastruktur etc.) und (ii) geschätzte Schäden an Hochwasser-Schutzbauten der Bundeswasserbauverwaltung.

- Sachschäden (gesamt) in monetären Einheiten: Die groben monetären Schätzungen zum Gesamtschaden werden nicht in das CESARE-Zielschema übernommen, da ein hohes Risiko von Überschneidung und daraus resultierenden Mehrfachnennungen mit den KatFonds-Datensätzen besteht.
- Schäden an HW-Schutzbauten: Die geschätzten Schäden an Hochwasser-Schutzbauten der Bundeswasserbauverwaltung werden in das CESARE-Zielschema zum Schadausmaß übernommen. Das Harmonisierungspotential ist grundsätzlich **hoch** und die Harmonisierung kann automatisiert erfolgen. MAXO-Werte zum Schadausmaß sind bereits im Original vorhanden und werden in das CESARE-Zielschema übernommen.

#### VIOLA

Im VIOLA-Datensatz finden sich, basierend auf Medienberichten, Angaben zum Ausmaß von (i) Personenschäden, (ii) Sachschäden in physischen Einheiten, (iii) Sachschäden in monetären Einheiten.

- Personenschäden: Angaben zum Ausmaß von Personenschäden werden in das CESARE-Zielschema übernommen. Das Harmonisierungspotential ist **hoch** und die Harmonisierung kann automatisiert erfolgen. Da es sich um Angaben aus Medienberichten handelt, wird als MAXO-Wert ein „A“ (Annahme/Schätzung) vergeben.

- Sachschäden in physischen Einheiten: Informationen zum Ausmaß von Sachschäden in physischen Einheiten werden derzeit, wie erwähnt, nicht vom CESARE-Zielschema erfasst.
- Sachschäden in monetären Einheiten: Die auf Medienberichten beruhenden und daher i. d. R. groben und vorläufigen Schadensschätzungen in monetären werden nicht in das CESARE-Zielschema übernommen, da ein hohes Risiko von Überschneidung und daraus resultierenden Mehrfachnennungen mit den KatFonds-Datensätzen besteht.

#### Einsatzdaten Landesfeuerwehr NÖ

Der Einsatzdatensatz des Landesfeuerwehrverbandes enthält keine monetären Angaben zu den Einsatzkosten oder dem Schadausmaß. Im Rahmen von CESARE wurde angedacht, die Einsatzkosten auf Basis des Original-Merkmals „anzahl\_ressourcen“ (Anzahl der am Einsatz beteiligten Ressourcen) abzuschätzen. Laut Auskunft des Datenbereitstellers handelt es sich dabei aber um eine Mischung aus Feuerwehren und Fahrzeugen und nicht unmittelbar um die Anzahl der am Einsatz beteiligten Personen. Eine seriöse Abschätzung der Einsatzkosten ist daher auf Basis der Daten nicht möglich.

#### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Der Datensatz „KatFonds NÖ Privat“ enthält monetäre Informationen über das Ausmaß von Katastrophenschäden am Vermögen physischer und juristischer Personen in Niederösterreich durch im Katastrophenfondsgesetz 1996 (KatFG 1996) gelistete Naturgefahren. Es handelt sich dabei um die von einer Schadenserhebungskommission geschätzten und in weiterer Folge als beihilfenfähig anerkannten Wiederherstellungskosten (Zeitwert), abzüglich etwaiger Versicherungsleistungen (siehe auch CESARE-Deliverable D4.1). Bei betrieblichen Einrichtungen werden die Wiederherstellungskosten exklusive Umsatzsteuer, ansonsten inklusive Umsatzsteuer angegeben.

In Bezug auf den Umfang der berücksichtigen Schadinformationen ist festzuhalten, dass etwaige unversicherte Schäden unter 1.000 Euro im Datensatz unterrepräsentiert sind, da der Gesamtschaden (abzüglich Versicherungsleistungen) grundsätzlich 1.000 Euro überschreiten muss, um beihilfenfähig zu sein. Nicht erfasst sind außerdem (vgl. Amt der Niederösterreichischen Landesregierung, 2005):

- Versicherte Schäden
- Schäden an landwirtschaftlichen Kulturen, die grundsätzlich versicherbar gewesen wären
- Schäden an Fischbeständen in Fließgewässern
- Schäden an Objekten, für die erforderliche behördliche Bewilligungen nicht vorliegen

Voraussetzung für eine Beihilfe ist eine spürbare materielle Belastung. Im Falle juristischer Personen muss eine Existenzgefährdung vorliegen. Größere Unternehmen sind daher selten vom Datensatz „KatFonds NÖ Privat“ erfasst.

Derzeit liegt uns der Datensatz und damit das Schadausmaß nur in Jahresauflösung, d. h. für jeden Schadeintrag ist nur das Jahr des Schadeintritts bekannt. Dadurch können die Informationen zum Schadausmaß aus dem Datensatz „KatFonds NÖ Privat“ nur für etwaige CESARE-Auswertungen auf Jahresbasis berücksichtigt werden. Zudem gilt, wie oben erwähnt, dass u. a. kleine Schäden sowie Schäden an größeren Unternehmen im Datensatz

unterrepräsentiert sind, was das Harmonisierungspotential zusätzlich etwas schmälert. Insgesamt wird das Harmonisierungspotential in Richtung Zielschema daher als **mittelniedrig** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Der zum Schadausmaß gehörende MAXO-Wert wird grundsätzlich auf „A“ (Annahme/Schätzung) gesetzt, da es sich i. d. R um Schätzungen der Wiederherstellungskosten handelt. Eine Ausnahme bilden jedoch die Original-Schadobjekte „Wegeschäden“, „Flussbauten“, „Wasserversorgungsanlagen“, „Wasserkraftwerke“ und „Teichanlagen“ sowie Schäden durch Erdbeben. Für sie muss zur endgültigen Festlegung der Schadenssumme eine Abrechnung der Schadenbehebung vorgelegt werden. In diesen Fällen wird der MAXO-Wert daher auf „M“ (Messung/Feststellung) gesetzt.

Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Der Datensatz „KatFonds NÖ Gemeinden“ enthält monetäre Informationen über das Ausmaß von Katastrophenschäden im Vermögen von Gemeinden in Niederösterreich durch im KatFG 1996 gelistete Naturgefahren. Das Schadausmaß ist im Datensatz mit zwei Merkmalen erfasst: die geschätzten Kosten und die Summe der nachgewiesenen Kosten. Bei den geschätzten Kosten handelt es sich um die von Sachverständigen bzw. einer Schadenkommission geschätzten Wiederherstellungskosten (Zeitwert). Bei der Summe der nachgewiesenen Kosten handelt es sich hingegen um die anhand von Rechnungen nachgewiesenen Wiederherstellungskosten abzüglich etwaiger Versicherungsleistungen (siehe auch CESARE-Deliverable D4.1). Handelt es sich beim Geschädigten um einen Betrieb mit marktwirtschaftlicher Tätigkeit (z.B. Wasser, Kanal, Kindergarten), beinhaltet der geschätzte bzw. ausgewiesene Betrag keine Umsatzsteuer, andernfalls schon.

Zwar existiert keine ziffernmäßig festgelegte Mindestschadenhöhe um als beihilfenfähig zu gelten, jedoch hängt es zum Teil von dem/der Bürgermeister:in ab, ob auch schon bei kleineren Schäden um Unterstützung angesucht und die Schäden damit beim Land gemeldet werden. Niedrige Schadenhöhen könnten somit unterrepräsentiert sein. Das Harmonisierungspotential wird daher als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Wo immer nachgewiesene Kosten vorhanden sind, werden diese in das CESARE-Zielschema (Merkmal „Schadenssumme“) übernommen und der zugehörige MAXO-Wert wird auf „M“ (Messung/Feststellung) gesetzt. Sind für einen Eintrag keine nachgewiesenen, aber geschätzte Kosten verfügbar, werden diese in das CESARE-Zielschema (Merkmal „Schadenssumme“) übernommen und für den dazugehörige MAXO-Wert wird ein „A“ (Annahme/Schätzung) vergeben. In diesem Fall ist zwar kein beihilfenfähiger, aber dennoch ein Schaden entstanden, dessen geschätzte Höhe in das CESARE-Zielschema aufgenommen wird.

Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Der Datensatz „KatFonds Stmk Privat“ enthält monetäre Informationen über das Ausmaß von Katastrophenschäden am Vermögen physischer und juristischer Personen in der Steiermark durch im Katastrophenschutzgesetz 1996 (KatFG 1996) gelistete Naturgefahren. Es handelt sich dabei um die von einer/m Amtssachverständigen oder einer/m allgemein beeideten und gerichtlich zertifizierten Sachverständigen geschätzten und in weiterer Folge als beihilfenfähig anerkannten Wiederherstellungskosten (Zeitwert), abzüglich etwaiger Versicherungsleistungen (siehe auch CESARE-Deliverable D4.1). Im Falle von Privatpersonen inkludiert die Schadenhöhe die Umsatzsteuer, im Falle vorsteuerabzugsberechtigter Betriebe ist hingegen keine Umsatzsteuer berücksichtigt.

In Bezug auf den Umfang der berücksichtigten Schadinformationen ist festzuhalten, dass etwaige unversicherte Schäden unter 1.000 Euro im Datensatz unterrepräsentiert sind, da der Gesamtschaden (abzüglich Versicherungsleistungen) 1.000 Euro überschreiten muss, um beihilfenfähig zu sein. Nicht erfasst sind außerdem u. a. (vgl. Amt der Steiermärkischen Landesregierung, 2018):

- Versicherte Schäden
- Schäden durch Hagel, Vermurung, Schneedruck und Bergsturz bei landwirtschaftlichen und gewerblichen Betrieben mit Ausnahme einer „De-minimis“-Beihilfe. Die Ausnahme gilt allerdings nicht für Hagelschäden bei landwirtschaftlichen Betrieben.
- Schäden an privaten Kraftfahrzeugen und privaten Wohnwägen, E-Bike etc.
- Schäden an Luxusgegenständen, Hobbygegenständen und -ausrüstungen etc.
- Nässe-, Dürre- und Frostschäden an landwirtschaftlichen Kulturen sowie Abschwemmschäden, die naturbedingt bei jedem stärkeren Regen auftreten können oder durch pflanzenbauliche Maßnahmen verhinderbar sind
- Private Teichanlagen und deren Besitz
- Fische in Fließgewässern
- Schäden an Glashäusern und Folientunnels in landwirtschaftlichen Betrieben
- Fallwild
- Ufereinrisse
- Soforthilfemaßnahmen (sind im Datensatz „*KatFonds Stmk Land*“ enthalten)

Voraussetzung für eine Beihilfe ist eine spürbare materielle Belastung. Im Falle juristischer Personen muss eine Existenzgefährdung vorliegen. Größere Unternehmen sind daher selten vom Datensatz „*KatFonds Stmk Privat*“ erfasst.

Da kleine Schäden sowie Schäden an größeren Unternehmen im Datensatz unterrepräsentiert sind und einige Schadkategorien, da nicht beihilfenfähig, ebenfalls nicht erfasst sind (z. B. private Kraftfahrzeuge, private Wohnwägen, private Teichanlagen etc.) wird das Harmonisierungspotential in Bezug auf das Zielschema als **mittel** eingestuft. Nachdem es sich beim Schadausmaß durchwegs um Schätzungen und nicht um nachgewiesene Kosten handelt, wird der dazugehörige MAXO-Wert auf „A“ (Annahme/Schätzung) gesetzt.

*Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)*

Der Datensatz „*KatFonds Stmk Land*“ enthält monetäre Informationen über das Ausmaß von Katastrophenschäden am steirischen Landesvermögen durch im Katastrophenfondsgesetz 1996 (KatFG 1996) gelistete Naturgefahren sowie Kosten durch Soforthilfemaßnahmen (siehe auch CESARE-Deliverable D4.1). Das angeführte Schadausmaß in Euro entspricht den geschätzten oder durch Rechnungen belegten Wiederherstellungskosten bzw. im Falle von Soforthilfemaßnahmen, die der Schadensvorbeugung dienen, den mit Rechnungen belegten Kosten dieser Maßnahmen. Teilweise sind auch Schadeinträge enthalten, für die nur das Monat oder das Jahr des Scha-

deintritts bekannt ist. Diese Einträge können nur für etwaige CESARE-Auswertungen auf Monats- oder Jahresbasis berücksichtigt werden. Das Harmonisierungspotential wird daher als **mittelhoch** eingestuft. Die Harmonisierung kann automatisiert erfolgen. Im Falle nachgewiesener Wiederherstellungs- oder Einsatzkosten wird der dazugehörige MAXO-Wert auf „M“ (Messung/Feststellung) gesetzt, im Falle geschätzter Kosten auf „A“ (Annahme/Schätzung). Ebenfalls ein „A“ vergeben wird für Schadeinträge, die im Zuge des Harmonisierungsprozesses aufgeteilt werden, nachdem im Original mehrere Schadereignisse unterschiedlicher Zeitpunkte zusammengefasst sind. Nicht in das CESARE-Zielschema übernommen werden etwaige dokumentierte Schadausmaße, die sich auf Vermögenswerte des Bundes beziehen (z. B. im Bereich der Wildbach- und Lawinenverbauung oder des Schutzwasserbaus).

### Verbleibende Herausforderungen im Bereich Schadausmaß

Zum Teil finden sich in den Original-Datensätzen unterschiedliche Kriterien, ab wann ein Schaden berücksichtigt wird. Das trifft insbesondere auf die KatFonds-Datensätze zu, wo es je nach Bundesland unterschiedliche Regelungen geben kann, welche Schäden beihilfenfähig sind und daher dokumentiert werden (inkl. vs. exkl. KFZ; inkl. vs. exkl. Teichanlagen; inkl. vs. exkl. Fischbestände etc.). Eine vollständige Harmonisierung gestaltet sich hier als schwierig. Zum Teil ist das betroffene Schadobjekt in dem Bundesland, wo es berücksichtigt wird, nicht als eigene Schadenkategorie ausgewiesen, sondern befindet sich in einem Aggregat. Dort, wo das nicht der Fall ist, liegen mit nur einem Bundesland derzeit oftmals noch zu wenig Daten vor, um für das andere Bundesland eine halbwegs robuste nachträgliche Schätzung für nicht berücksichtigte Objekte zu ermöglichen. Wir beschränken uns daher vorläufig auf die Bereitstellung der entsprechenden Metainformation zum Umfang der berücksichtigten bzw. dokumentierten Schäden.

Dass versicherte Schäden in den KatFonds-Datensätzen nicht berücksichtigt sind, stellt in Bezug auf die Harmonisierung grundsätzlich kein Problem dar. Bei Daten zu versicherten Schäden, die für CESARE derzeit aber nicht vorliegen, handelt es sich somit um komplementäre Information ohne jegliche Überschneidungsgefahr mit den KatFonds-Datensätzen.

## OBJEKTBEZITZER:IN (BEI SACHSCHÄDEN)

### Definition des CESARE Zielschemas

Basierend auf den Empfehlungen des Joint Research Centers (JRC, 2015) sieht das CESARE-Zielschema im Falle von Sachschäden auch die Erfassung des Objektbesitzers bzw. der Objektbesitzerin über folgendes Merkmal vor (siehe auch CESARE-Deliverable D4.1):

- Objektbesitzer (enumeration)

Aufbauend auf vorhandenen Informationen in den zur Verfügung stehenden Datensätzen wird im Merkmal „Objektbesitzer“ zwischen „Physische/juristische Person“, „Gebietskörperschaft - Gemeinde“, „Gebietskörperschaft - Land“ und „Gebietskörperschaft - Bund“ bzw. „Gebietskörperschaft“ (falls die konkrete Verwaltungsebene nicht bekannt ist oder es sich um mehrere handelt) unterschieden.

## Harmonisierungspotential und Umsetzung je Datensatz

### GEORIOS

Es werden keine Sachschäden aus dem GEORIOS-Datensatz in das CESARE-Zielschema überführt.

### WLK

Aus dem WLK-Datensatz werden Schätzungen zu Wiederherstellungskosten von Schäden an WLV-Verbauungen in das CESARE-Zielschema übernommen. Der Objektbesitz wird im CESARE-Zielschema der Kategorie „Gebietskörperschaft“ zugeordnet. Da es sich auch um Konkurrenzanlagen handeln kann, erfolgt keine genauere Spezifizierung der Verwaltungsebene. Das Harmonisierungspotential ist **hoch**.

### HWFDB

Aus dem HWFDB-Datensatz werden Schätzungen zu Wiederherstellungskosten von Schäden an HW-Schutzbauten in das CESARE-Zielschema übernommen. Der Objektbesitz wird im CESARE-Zielschema der Kategorie „Gebietskörperschaft“ zugeordnet. Da es sich auch um Konkurrenzanlagen handeln kann, erfolgt keine genauere Spezifizierung der Verwaltungsebene. Das Harmonisierungspotential ist **hoch**.

### VIOLA

Es werden keine Sachschäden aus dem VIOLA-Datensatz in das CESARE-Zielschema überführt.

### Einsatzdaten Landesfeuerwehr NÖ

Es werden keine Sachschäden aus dem Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich in das CESARE-Zielschema überführt.

### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Der Datensatz „KatFonds NÖ Privat“ enthält ausschließlich Informationen zu Katastrophenschäden am Vermögen physischer und juristischer Personen. Der Objektbesitz ist im CESARE-Zielschema daher der Kategorie „Physische/juristische Person“ zuzuordnen. Das Harmonisierungspotential ist **hoch**.

### Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Der Datensatz „KatFonds NÖ Gemeinden“ enthält ausschließlich Informationen zu Katastrophenschäden am Vermögen von Gemeinden. Der Objektbesitz ist im CESARE-Zielschema daher der Kategorie „Gebietskörperschaft – Gemeinde“ zuzuordnen. Das Harmonisierungspotential ist **hoch**.

### Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Der Datensatz „KatFonds Stmk Privat“ enthält ausschließlich Informationen zu Katastrophenschäden am Vermögen physischer und juristischer Personen. Der Objektbesitz ist im CESARE-Zielschema daher der Kategorie „Physische/juristische Person“ zuzuordnen. Das Harmonisierungspotential ist **hoch**.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Der Datensatz „KatFonds Stmk Land“ enthält neben Informationen zu Katastrophenschäden am Landesvermögen zum Teil auch Informationen zu Katastrophenschäden am Vermögen physischer und juristischer Personen (Soforthilfemaßnahmen), am Gemeindevermögen (Soforthilfemaßnahmen sowie Konkurrenzanalgen im Bereich Wildbach- und Lawinenverbauung sowie Schutzwasserbau) und am Bundesvermögen (Konkurrenzanalgen im Bereich Wildbach- und Lawinenverbauung sowie Schutzwasserbau). Informationen zu Schäden bzw. Wiederherstellungskosten am Bundesvermögen fließen allerdings, wie erwähnt, nicht in das CESARE-Zielschema ein. Bis auf wenige Ausnahmen geht die Information über den Objektbesitz aus den Originaldaten hervor. Das Harmonisierungspotential kann somit als **hoch** eingestuft werden und die Harmonisierung automatisiert erfolgen.

**Verbleibende Herausforderungen im Bereich Objektbesitz**

Keine

**SCHADENTRÄGER:IN**

**Definition des CESARE Zielschemas**

Basierend auf den Empfehlungen des Joint Research Centers (JRC, 2015) sieht das CESARE-Zielschema die Zuordnung bzw. Aufteilung der erfassten Wiederherstellungs- und Einsatzkosten auf die jeweiligen Schadenträger:innen vor (siehe auch CESARE-Deliverable D4.1). Nachdem sich die Finanzierung der Wiederherstellungs- und Einsatzkosten für ein und denselben Schadeintrag auf unterschiedliche Schadenträger:innen aufteilen kann, wird diese Information in einer separaten Tabelle geführt, die über einen eindeutigen Identifier mit der eigentlichen Schadentabelle verbunden ist. Neben diesem Identifier umfasst das Zielschema folgende Merkmale:

- Schadenssumme (numeric)
- Schadentraeger (enumeration)
- Schadentraeger\_MAXO (enumeration: {M, A, X, O})

Das Merkmal „Schadenssumme“ entspricht jenem in der eigentlichen Schadentabelle, abgesehen davon, dass sich die Schadenssumme auf mehrere Schadenträger:innen verteilen kann. Das Merkmal „Schadenträger“ gibt an, wer letztendlich für die im Merkmal „Schadenssumme“ angeführten Wiederherstellungs- und Einsatzkosten aufgekommen ist. Dabei wird zwischen „Physische/juristische Person“, „Gebietskörperschaft - Gemeinde“, „Gebietskörperschaft - Land“, „Gebietskörperschaft - Bund“, „Gebietskörperschaft“ (falls die konkrete Verwaltungsebene nicht bekannt ist oder es sich um mehrere Ebenen handelt) und „Versicherung“ unterschieden. Das zugehörige MAXO-Merkmal informiert über den Grad der Zuverlässigkeit des Schadausmaßes, das dem jeweiligen „Schadenträger“ zugeordnet ist.

## Harmonisierungspotential und Umsetzung je Datensatz

### GEORIOS

Es werden keine Wiederherstellungs- oder Einsatzkosten aus dem GEORIOS-Datensatz in das CESARE-Zielschema überführt.

### WLK

Aus dem WLK-Datensatz werden Schätzungen zu den vom Bund getragenen Wiederherstellungskosten beschädigter WLV-Verbauungen in das CESARE-Zielschema übernommen. Diese werden im CESARE-Zielschema daher der Schadenträger-Kategorie „Gebietskörperschaft – Bund“ zugeordnet. Das Harmonisierungspotential ist hoch.

### HWFDB

Aus dem HWFDB-Datensatz werden Schätzungen zu den vom Bund getragenen Wiederherstellungskosten beschädigter HW-Schutzbauten der Bundeswasserbauverwaltung in das CESARE-Zielschema übernommen. Diese werden im CESARE-Zielschema daher der Schadenträger-Kategorie „Gebietskörperschaft – Bund“ zugeordnet. Das Harmonisierungspotential ist hoch.

### VIOLA

Es werden keine Wiederherstellungs- oder Einsatzkosten aus dem VIOLA-Datensatz in das CESARE-Zielschema überführt.

### Einsatzdaten Landesfeuerwehr NÖ

Es werden keine Wiederherstellungs- oder Einsatzkosten aus dem Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich in das CESARE-Zielschema überführt.

### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Bei außergewöhnlichen Schäden im Vermögen physischer und juristischer Personen durch im KatFG 1996 gelistete Naturgefahren gewährt das Land den geschädigten Personen Beihilfen, die je nach Art des geschädigten Objekts in der Regel zwischen bis zu 20 % und bis zu 70 % der als beihilfenfähig anerkannten Schäden ausmachen können (Amt der Niederösterreichischen Landesregierung, 2005). Im Datensatz „KatFonds NÖ Privat“ ist die Beihilfenhöhe unter dem Merkmal „Beihilfe“ dokumentiert. Von dieser vom Land gewährten Beihilfe werden in der Regel 60 % durch den Katastrophenfonds des Bundes refundiert. Auf Basis dieser Informationen lässt sich für jeden Schadeintrag ableiten, welcher Teil des Schadens von der geschädigten privaten bzw. juristischen Person, welcher Teil vom Land und welcher Teil vom Bund getragen wurde. Das Harmonisierungspotential kann als **hoch** eingestuft werden und die Harmonisierung automatisiert erfolgen. Der MAXO-Wert für den Landes- und Bundesanteil wird auf „M“ (Messung/Feststellung) gesetzt, jener für den von der geschädigten Person getragene Anteil orientiert sich am MAXO-Wert der Schadenssumme in der eigentlichen Schadentabelle. Der Datensatz enthält 10 Einträge, wo das Schadenausmaß mit 0 beziffert ist, die Beihilfenhöhe jedoch einen positiven Wert auf-

weist sowie einen Eintrag, wo die Beihilfenhöhe über dem Schadenausmaß liegt. Für diese Fälle wird die Beihilfenhöhe auf das angegebene Schadenausmaß revidiert und der zum Merkmal „Schadenträger“ gehörende MAXO-Wert auf „A“ (Annahme/Schätzung) gesetzt.

Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Für die Beseitigung von außergewöhnlichen Schäden im Vermögen von Gemeinden durch im KatFG 1996 gelistete Naturgefahren können Katastrophenfondsmittel des Bundes als Beihilfe gewährt werden. Diese betragen in der Regel 50 % der nachgewiesenen Kosten bzw. im Falle von Forstschäden 18 % der nachgewiesenen Kosten. Grundsätzlich kann die konkrete Beihilfenhöhe dem Original-Merkmal „Summe Anweisungen“ entnommen werden. Bei älteren Schadeinträgen können die dortigen Summen allerdings von 50 % bzw. 18 % der nachgewiesenen Kosten abweichen. Früher gab es die Möglichkeit, bei besonderen Schadereignissen zunächst ohne Vorlage von Rechnungsnachweisen einen Vorschuss aus dem Katastrophenfonds in der Höhe von 20 % der geschätzten Kosten zu beantragen. Bei ausbleibender Nachreichung der Rechnungsnachweise wurde der Vorschuss bei späteren Förderungen gegengerechnet. Im Rahmen des CESARE-Harmonisierungsprozesses erfolgt daher je Schadeintrag eine Nachberechnung der dazugehörigen Beihilfenhöhe des Bundes unter Anwendung der 50 % bzw. 18 %-Regel auf die Summe der nachgewiesenen Kosten. Der Rest der nachgewiesenen Kosten ist von den Gemeinden selbst zu tragen. Das Harmonisierungspotential ist **hoch** und die Harmonisierung kann automatisiert erfolgen. Der MAXO-Wert für das Merkmal „Schadenträger“ orientiert sich am MAXO-Wert der Schadensumme in der eigentlichen Schadentabelle.

Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Bei außergewöhnlichen Schäden im Vermögen physischer und juristischer Personen durch im KatFG 1996 gelistete Naturgefahren gewährt das Land den geschädigten Personen Beihilfen, die je nach Art des geschädigten Objekts in der Regel zwischen 30 % und 50 % der als beihilfenfähig anerkannten Schäden ausmachen können (Amt der Steiermärkischen Landesregierung, 2018). Im Datensatz „KatFonds Stmk Privat“ ist die Beihilfenhöhe unter dem Merkmal „BEIHILFE“ dokumentiert. Von dieser vom Land gewährten Beihilfe werden in der Regel 60 % durch den Katastrophenfonds des Bundes refundiert. Auf Basis dieser Informationen lässt sich für jeden Schadeintrag ableiten, welcher Teil des Schadens von der geschädigten privaten bzw. juristischen Person, welcher Teil vom Land und welcher Teil vom Bund getragen wurde. Das Harmonisierungspotential kann als **hoch** eingestuft werden und die Harmonisierung automatisiert erfolgen. Der MAXO-Wert für den Landes- und Bundesanteil wird auf „M“ (Messung/Feststellung) gesetzt, jener für den von der geschädigten Person getragene Anteil orientiert sich am MAXO-Wert der Schadensumme in der eigentlichen Schadentabelle.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Für die Beseitigung von außergewöhnlichen Schäden im Landesvermögen durch im KatFG 1996 gelistete Naturgefahren können Katastrophenfondsmittel des Bundes als Beihilfe gewährt werden. In den zur Verfügung gestellten Daten – hauptsächlich pdf-Scans – sind die Beihilfenhöhen zwar nicht pro Schadeintrag ausgewiesen, es finden sich pro Jahr jedoch Übersichten mit der Summe der gemeldeten Schäden pro Fachabteilung und den erhaltenen Katastrophenfondsmitteln. Aus diesen Übersichten und dem KatFG 1996 lassen sich folgende Regelungen ableiten:

- Abgesehen von den nachfolgend gelisteten Ausnahmen werden 50 % der Wiederherstellungskosten für Schäden im Landesvermögen vom Bund über den Katastrophenfonds getragen
- Bei Landesstraßen B trägt der Bund 50 % jener Wiederherstellungskosten, die den Sockelbetrag von 2.520.000 Euro im Jahr übersteigen
- Bei Schäden im Bereich der Landesbahnen trägt der Bund ab 2013 18 % (davor 50 %) der Wiederherstellungskosten
- Bei Schäden im Bereich der Landesforste sowie land-/forstwirtschaftlichen Fachschulen trägt der Bund ab 2016 18 % bzw. im Fall von Hagelschäden 0 % (davor 50 %) der Wiederherstellungskosten
- Bei Schäden im Vermögen privater und juristischer Personen, die vom Land mittels Sofortmaßnahmen behoben wurden, refundiert der Bund dem Land 60 % der Kosten (vgl. Amt der Steiermärkischen Landesregierung, 2013)
- Bei Schäden im Vermögen von Gemeinden, die vom Land mittels Sofortmaßnahmen behoben wurden, refundiert der Bund dem Land 50 % der Kosten (vgl. Amt der Steiermärkischen Landesregierung, 2013)
- Kosten für Soforthilfemaßnahmen, die nicht der Wiederherstellung eines beschädigten Objekts dienen, sondern reine Vorbeugemaßnahmen ohne verbundenem Schaden darstellen, sind zur Gänze vom Land zu tragen

Das Harmonisierungspotential ist als **hoch** einzustufen und kann über die oben genannten Regeln automatisiert erfolgen. Nachdem die Aufteilung der Wiederherstellungs- und Einsatzkosten auf die entsprechenden Schadenträger:innen für die einzelnen Schadeinträge komplett ex-post erfolgt, wird der MAXO-Wert zum Merkmal „Schadenträger“ auf durchgehend auf „A“ (Annahme/Schätzung) gesetzt.

Verbleibende Herausforderungen im Bereich Schadenträger:in

Keine

## DATENQUELLE

### Definition des CESARE Zielschemas

Ein weiteres Merkmal des CESARE-Zielschemas stellt die Datenquelle dar:

- Datenquelle (character)

Anhand des Merkmals wird für jeden Ereignis- und/oder Schadeintrag die Datenquelle mit Namen des Datensatzes, Namen der bereitstellenden Institution bzw. Organisation und Datum der Bereitstellung dokumentiert. Das Merkmal wird auch dazu verwendet, Teile von Datensätzen, die einen Strukturbruch beinhalten, separat auszuweisen und so ein rasches Zu- und Wegschalten des Subdatensatzes zu ermöglichen.

## Harmonisierungspotential und Umsetzung je Datensatz

### GEORIOS

Der GEORIOS-Datensatz ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „GEORIOS, GBA, Stand 25.03.2021“.

### WLK

Der GEORIOS-Datensatz ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „WLK, WLW, Stand 23.11.2020“.

### HWFDB

Der GEORIOS-Datensatz ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „HWFDB, Bundeswasserbauverwaltung, Stand 1.9.2020“.

### VIOLA

Der VIOLA-Datensatz ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „VIOLA, ZAMG, Stand 26.08.2021“.

### Einsatzdaten Landesfeuerwehr NÖ

Der Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „Einsatzdaten, Niederösterreichischer Landesfeuerwehrverband, Stand 7.6.2021“.

### Landesverwaltung NÖ – Katastrophenschäden am Privatvermögen (KatFonds NÖ Privat)

Der Datensatz „KatFonds NÖ Privat“ weist in Bezug auf die dokumentierten Schäden in der Landwirtschaft einen Strukturbruch auf (siehe auch CESARE-Deliverable D4.1). Seit Mitte 2016 hat sich damit der Umfang der erfassten Schäden an landwirtschaftlichen Kulturen deutlich verringert. Für den Datensatz werden daher zwei unterschiedliche Quellnamen vergeben:

- „Katastrophenschäden Privatvermögen, Land Noe, Stand 30.03.2021“ für alle Einträge, die sich nicht auf Schäden an landwirtschaftlichen Kulturen beziehen
- „Katastrophenschäden Privatvermögen (Landwirtschaftl. Kulturen), Land Noe, Stand 30.03.2021“ für alle Einträge, die sich auf Schäden an landwirtschaftlichen Kulturen beziehen

Auf diese Weise kann jener Teil des Datensatzes, der einen Strukturbruch aufweist, bei Bedarf auf einfache Art und Weise ausgeblendet werden.

### Landesverwaltung NÖ – Katastrophenschäden am Gemeindevermögen (KatFonds NÖ Gemeinden)

Der Datensatz „KatFonds NÖ Gemeinden“ ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „Katastrophenschäden Gemeindevermögen, Land Noe, Stand 17.12.2020“.

Landesverwaltung Steiermark – Katastrophenschäden am Privatvermögen (KatFonds Stmk Privat)

Der Datensatz „KatFonds Stmk Privat“ ist in der derzeitigen Version mit folgendem Quellnamen ausgewiesen: „Katastrophenschäden Privatvermögen, Land Stmk, Stand 15.03.2021“.

Landesverwaltung Steiermark – Katastrophenschäden am Landesvermögen (KatFonds Stmk Land)

Der Datensatz „KatFonds Stmk Land“ besteht aus zwei Teil-Datensätzen: der erste dokumentiert die Einsatz- und Wiederherstellungskosten durch Soforthilfe- und Folgemaßnahmen im Katastrophenfall, der zweite die Wiederherstellungskosten im Nicht-Katastrophenfall. Letzterer weist einen Strukturbruch auf: Wiederherstellungskosten für Schäden im Bereich der Wildbach- und Lawinerverbauung sowie des Schutzwasserbaus sind dort nur bis einschließlich 2013 erfasst (siehe auch CESARE-Deliverable D4.1). Für den zweiten Teildatensatz werden daher zwei unterschiedliche Quellnamen vergeben. Damit sind insgesamt drei unterschiedliche Quellnamen in Verwendung:

- „Katastrophenschäden Soforthilfemaßnahmen, Land Stmk, Stand 06.04.2021“ für den ersten Teil-Datensatz
- „Katastrophenschäden Landesvermögen, Land Stmk, Stand 06.04.2021“ für alle Einträge im zweiten Teil-Datensatz, die sich nicht auf Schäden im Bereich Wildbach- und Lawinerverbauung oder Schutzwasserbau beziehen
- „Katastrophenschäden Landesvermögen (WLV und Schutzwasserbau), Land Stmk, Stand 06.04.2021“ für alle Einträge im zweiten Teil-Datensatz, die sich auf Schäden im Bereich Wildbach- und Lawinerverbauung oder Schutzwasserbau beziehen

**Verbleibende Herausforderungen im Bereich Verortung**

Keine

## EREIGNIS-ID

**Definition des CESARE Zielschemas**

Die für CESARE zur Verfügung stehenden Datensätze können grob in zwei Klassen unterteilt werden:

- Datensätze, in denen **Ereignisse** das Hauptelement darstellen. Dazu zählen GEORIOS, WLK, HWFDB und VIOLA. Bei der Erfassung eines neuen Ereignisses wird in diesen Datensätzen jeweils eine Ereignis-ID vergeben. Das „Ereignis“ bildet dann in der Regel die räumlich/zeitliche Bezugseinheit. Je nach Datensatz können dieser Bezugseinheit mehrere Phänomene und/oder Schäden zugeordnet sein.
- Datensätze, in denen **Schäden im weiteren Sinne** das Hauptelement darstellen. Dazu zählen die einzelnen „KatFonds“-Datensätze und der Einsatzdatensatz des Landesfeuerwehrverbandes Niederösterreich. Hier werden in der Regel für einzelne Schadenfälle und Einsätze IDs vergeben. Eine Zuordnung zu bestimmten Ereignissen findet üblicherweise nicht statt.

Im CESARE Zielschema stellen Ereignisse das Hauptelement dar und spannen die Bezugseinheit in Hinblick auf Zeit, Raum und Naturgefahr auf. Ihnen können in weiterer Folge Phänomene und Schäden zugeordnet werden. Im Zuge des Harmonisierungsprozesses gilt es daher zum einen, bestehende Ereignisse aus unterschiedlichen Quellen, in denen Ereignisse das Hauptelement darstellen, zusammenzuführen. Zum anderen sind die Schadenfälle und Einsätze aus den Datensätzen, in denen Schäden im weiteren Sinne das Hauptelement darstellen, existierenden oder neu zu definierenden Ereignissen zuzuordnen. Die Ereignisdefinition und Ereignis-ID-Vergabe sollen dabei möglichst automatisiert erfolgen. Im Zuge von CESARE wurde hierfür ein pilothafter Algorithmus entwickelt, den es in Zukunft noch weiter zu verfeinern gilt. Die Ereigniszusammenführung bzw. Ereignis-ID-Vergabe kann sich dabei abhängig von der betrachteten Naturgefahr unterscheiden, basiert im Kern aber jeweils auf dem zeitlichen Auftreten, der räumlichen Entfernung sowie der Art des Ereignisses bzw. Prozesses. Das Startdatum eines CESARE-Ereignisses wird über das gemeinsame Startdatum der einzelnen zusammengeführten Einträge festgelegt. Das Enddatum eines CESARE-Ereignisses definiert sich in der Regel aus dem spätesten Enddatum der zusammengefassten Einzeleinträge. Z. T wird bei der Festlegung des Enddatums allerdings auch abgeschnitten, da sich in manchen Datenquellen Unsicherheiten in der Datumsangabe durch lange Zeitspannen zwischen Start- und Enddatum ausdrücken (siehe z. B. KatFonds NÖ Gemeinden). In diesem Fall erfolgt eine entsprechende Kennzeichnung mittels MAXO-Code („A“).

Im Folgenden wird die Ausgestaltung des Algorithmus je Naturgefahr beschrieben, wobei neben den CESARE-Gefahren Hochwasser, Wind/Sturm und Massenbewegungen auch jene Gefahren mitberücksichtigt werden, die in den zur Verfügung gestellten Datensätzen darüber hinaus vorkommen. Die Zusammenführung der Ereignisse über die unterschiedlichen Datenquellen hinweg erfolgt basierend auf einer 7-stelligen Ereignis-ID, der ein „E“ vorangestellt ist:

- Event\_ID („E“ + 7-digit integer)

#### Algorithmus zur Ereignisidentifikation bzw. Ereignis-ID-Vergabe

Im Folgenden wird der CESARE-Algorithmus zur Ereignisidentifikation bzw. Ereignis-ID-Vergabe für jede Naturgefahr separat beschrieben. In Abbildung A-1 ist der gesamte Prozess in Form eines Entscheidungsbaums dargestellt.

##### Massenbewegungen & Schneelawinen

Bei Massenbewegungen und Schneelawinen erfolgt die Ereignisdefinition in CESARE möglichst kleinteilig, so wie das auch in den Originaldatensätzen GEORIOS und WLK der Fall ist. Sofern es der Originaldatensatz erlaubt, wird in CESARE jeder einzelne Massenbewegungsprozess und jede einzelne Schneelawine als separates Ereignis definiert. Handelt es sich um einen Datensatz, in dem Ereignisse das Hauptelement darstellen, werden die dort definierten Ereignisse demnach eins zu eins übernommen und erhalten eine CESARE-Ereignis-ID. Handelt es sich hingegen um einen KatFonds-Datensatz, werden alle Schadeinträge, die dasselbe Startdatum, dieselbe Gemeinde (bzw. dieselben Gemeinden, falls ein Schadeintrag mehrere Gemeinden umfasst) und dieselbe Schadenursache (Gefahr/Hazard auf Ebene 3) aufweisen, zu einem Ereignis zusammengefasst. Da die KatFonds-Datensätze auf Gemeindeebene vorliegen, lässt sich aus den verfügbaren Informationen nicht ableiten, ob sich mehrere

Schadeinträge mit demselben Datum und derselben Gemeinde auf ein einziges Ereignis oder mehrere Ereignisse beziehen.

#### Hochwasser

Im Falle der Naturgefahr „Hochwasser“ werden bekannte Großereignisse und ihre zeitliche und räumliche Ausdehnung (i. S. v. betroffenes Bundesland) vorab über die ZAMG-Unwetterchroniken, VIOLA und Internetrecherche identifiziert. Hochwasserereignisse und -schadeinträge aus den Originaldaten, die zeitlich und räumlich mit den vorab identifizierten Großereignissen zusammenfallen, werden diesen zugeordnet. Für die restlichen Hochwasserereignisse und -schadeinträge der unterschiedlichen Originalquellen erfolgt eine Gruppierung nach Startdatum und aneinander angrenzenden Gemeinden (Gemeindestand 2020). Hochwasserereignisse und -schadeinträge, die am selben Tag (Startdatum) und in räumlich aneinandergrenzenden Gemeinden verzeichnet wurden, werden zu einem CESARE-Ereignis zusammengefasst.

#### Wind/Sturm

Für die Naturgefahr „Wind/Sturm“ wird analog vorgegangen wie für die Naturgefahr „Hochwasser“. Ereignisse und Schadeinträge aus den Originaldaten, die zeitlich und räumlich mit vorab identifizierten Großereignissen der Naturgefahr „Wind/Sturm“ zusammenfallen, werden diesen zugeordnet. Für die restlichen Sturmereignisse und -schadeinträge der unterschiedlichen Originalquellen erfolgt eine Gruppierung nach Startdatum und aneinander angrenzenden Gemeinden.

#### Niederschlag (Gewitter/Starkregen/Dauerregen)

Für die Niederschlagskategorien „Gewitter“, „Starkregen“ und „Dauerregen“, die bei der Ereignis-ID-Vergabe als Synonyme behandelt werden, wird analog vorgegangen wie für die Naturgefahr „Hochwasser“. Ereignisse und Schadeinträge aus den Originaldaten, die zeitlich und räumlich mit vorab identifizierten Großereignissen im Bereich Gewitter/Starkregen/Dauerregen zusammenfallen, werden diesen zugeordnet. Für die restlichen Ereignisse und Schadeinträge der unterschiedlichen Originalquellen erfolgt eine Gruppierung nach Startdatum und aneinander angrenzenden Gemeinden.

#### Niederschlag (Hagel)

Ereignisse und Schadeinträge der Niederschlagskategorie „Hagel“ werden nach Startdatum und aneinander angrenzenden Gemeinden gruppiert, d. h. Hagelereignisse und -schadeinträge, die am selben Tag (Startdatum) und in räumlich aneinandergrenzenden Gemeinden verzeichnet wurden, werden zu einem Ereignis zusammengefasst.

#### Niederschlag (Schneefall bzw. Schneedruck)

Die Niederschlagskategorie „Schneefall“ kommt einerseits im VIOLA-Datensatz vor, wo entsprechende Ereignisse dokumentiert sind, andererseits aber auch in den KatFonds-Datensätzen, die unter der Originalbezeichnung „Schneedruck“ entsprechende Schadeinträge erfassen. Der CESARE-Algorithmus übernimmt die in VIOLA dokumentierten Ereignisse eins zu eins und vergibt lediglich eine eigenständige CESARE-Ereignis-ID. Für die Schadeinträge in den KatFonds-Datensätzen wird hingegen untersucht, ob deren Startdatum in die Zeitspanne eines

VIOLA-Ereignisses fällt und sich die betroffenen Gemeinden überschneiden. Ist das der Fall, wird der jeweilige Schadeintrag dem entsprechenden Ereignis zugeordnet. Für die restlichen KatFonds-Einträge erfolgt die Gruppierung der Schadfälle nach Startdatum und Bundesland, wobei direkt aneinander anschließende Tage (Startdatum) zu einem Ereignis zusammengefasst werden.

### Erdbeben

Die Naturgefahr „Erdbeben“ kommt in den zur Verfügung gestellten Daten nur in Datensätzen vor, in denen Schäden im weiteren Sinne das Hauptelement darstellen (KatFonds-Datensätze und Einsatzdaten der Feuerwehr). Der CESARE-Algorithmus gruppiert Einsätze und Schäden innerhalb eines Bundeslandes nach dem Startdatum, wobei Einsätze und/oder Schäden mit unmittelbar zusammenhängenden Starttagen zu einem Ereignis zusammengefasst werden.

### Waldbrand

Die Gefahr „Waldbrand“ kommt in den zur Verfügung gestellten Daten lediglich in den Einsatzdaten des Landesfeuerwehrverbandes Niederösterreich vor. Einsätze an aufeinander folgenden Tagen und in aneinander angrenzenden Gemeinden werden vom CESARE-Algorithmus zu einem Ereignis zusammengefasst.

### Fehlende Naturgefahr oder fehlendes Datum

Ereignisse oder Schadeinträge ohne Angaben zum konkreten Startdatum oder zur konkreten Naturgefahr werden jeweils als einzelne, separate Ereignisse definiert und erhalten eine eigene CESARE-Ereignis-ID.

### **Verbleibende Herausforderungen im Bereich Ereignisidentifizierung bzw. Ereignis-ID-Vergabe**

Eine Katastrophe kann in der einen Datenquelle als mehrere Einzelereignisse registriert sein und in einer anderen als ein großes Ereignis. Hinzu kommen Unschärfen in Bezug auf die Datums-/Zeitangabe, Verortung und Naturgefahr. Dies führt dazu, dass auch nach erfolgter Harmonisierung im Sinne von Ereigniszusammenführung sich überschneidende Ereignisse oder Mehrfachnennungen existieren können. Eine Auswertung des zusammengeführten CESARE-Datensatzes in Bezug auf die Anzahl der Ereignisse ist daher nur bedingt sinnvoll und aussagekräftig. Wie im Unterabschnitt zum Schadausmaß beschrieben, wird im Zuge der Harmonisierung aber darauf geachtet, das Risiko von Mehrfachzählungen in Hinblick auf das Schadausmaß möglichst gering zu halten.

In Zukunft gilt es auch, den Algorithmus zur Ereigniszusammenführung noch weiter zu optimieren. Derzeit gruppiert der Algorithmus basierend auf dem (Start-)Datum, der geografischen Nähe und der Überschneidung der Naturgefahrenart bzw. bestimmter Naturgefahrenarten. Noch nicht im automatisierten Algorithmus integriert sind hingegen meteorologische Daten. Ein Einbezug meteorologischer Daten bei der Ereignisidentifikation und -zusammenführung wurde versuchsweise bereits im Rahmen von CESARE in einem semi-automatisierten Prozess mit hohem manuellen Anteil erprobt. Vor allem für größere Ereignisse liefert er eine deutliche Verbesserung bei der Zusammenfassung zusammenhängender Einträge, ist aber auch entsprechend zeitaufwendig. Zukünftiges Ziel ist daher, meteorologische Daten in den automatisierten Algorithmus einzubauen und diesen weiter zu verfeinern.

## KOMPOSIT-ID

### Definition des CESARE Zielschemas

Neben der Ereignis-ID sieht das CESARE-Zielschema auch die Vergabe einer Komposit-ID vor. Die Vergabe von Ereignis-IDs erfolgt in CESARE immer innerhalb ein- und derselben Naturgefahren- bzw. Prozessart. Unterschiedliche Prozessarten können jedoch durch denselben Auslöser verbunden sein, beispielsweise im Falle eines Unwetters, das Muren auslöst, zu Hochwasser führt und Sturmschäden verursacht. Die Komposit-ID fasst Ereignisse zusammen, denen vermutlich derselbe Auslöser zugrunde liegt bzw. die aller Wahrscheinlichkeit nach zusammenhängen. Bei der Komposit-ID handelt es sich also um eine Verlinkung von Ereignis-IDs. Im Zuge von CESARE wurde hierfür ein pilothafter Algorithmus entwickelt. Er ähnelt stark dem Algorithmus zur Ereignis-ID-Vergabe mit dem Unterschied, das originale Ereignis- bzw. Schadeinträge über Naturgefahren- bzw. Prozesskategorien hinweg zusammengefasst werden. Wie auch die Ereignis-ID-Vergabe kann sich die Ereigniszusammenführung im Zuge der Komposit-ID-Vergabe abhängig von der betrachteten Naturgefahr unterscheiden. Im Folgenden wird die Ausgestaltung des Algorithmus näher beschrieben. Die Zusammenführung der Ereignisse über die unterschiedlichen Datenquellen hinweg erfolgt basierend auf einer 7-stelligen Ereignis-ID, der ein „C“ vorangestellt ist:

- Composit\_ID („C“ + 7-digit integer)

### Algorithmus zur Komposit-ID-Vergabe

Im Folgenden wird der CESARE-Algorithmus zur Komposit-ID-Vergabe beschrieben. In Abbildung A-2 ist der gesamte Prozess in Form eines Entscheidungsbaums dargestellt.

#### Schneelawine, Waldbrand, Erdbeben

Für die Naturgefahren „Schneelawine“, „Waldbrand“ und „Erdbeben“ erfolgt die Vergabe der Komposit-ID analog zur Vergabe der Ereignis-ID. Mit anderen Worten werden für diese Naturgefahren die mittels CESARE-Algorithmus definierten Ereignisse nicht weiter zusammengefasst. Jede CESARE-Ereignis-ID entspricht einer separaten CESARE-Komposit-ID.

#### Hochwasser, Massenbewegungen, Wind/Sturm, Niederschlag

Im Falle der Naturgefahren „Hochwasser“, „Massenbewegung“, „Wind/Sturm“ und „Niederschlag“ werden bekannte, sich aus unterschiedlichen Prozessen zusammensetzende Großereignisse und ihre zeitliche und räumliche Ausdehnung (i. S. v. betroffenen Bundesländern) vorab über die ZAMG-Unwetterchroniken, VIOLA und Internetrecherche identifiziert. Ereignisse und Schadeinträge aus den Originaldatensätzen, die einer der genannten Naturgefahren zugeordnet sind sowie zeitlich und räumlich mit den vorab identifizierten Großereignissen zusammenfallen, werden diesen zugeordnet. Für die restlichen Ereignisse und Schadeinträge der unterschiedlichen Originalquellen erfolgt eine Gruppierung nach Startdatum, aneinander angrenzenden Gemeinden (Gemeindestand 2020) und CESARE-Ereignis-ID. Das heißt, Ereignisse und Schadeinträge aus den Originaldatensätzen, die einer der genannten Naturgefahren zugeordnet sind, dasselbe Startdatum aufweisen und in räumlich

aneinandergrenzenden Gemeinden verzeichnet wurden, werden zu einem CESARE-Komposit zusammengefasst, wobei Einträge mit derselben CESARE-Ereignis-ID zwingend eine einheitliche Komposit-ID erhalten.

#### Fehlende Naturgefahr oder fehlendes Datum

Für Ereignisse oder Schadeinträge ohne Angaben zum konkreten Startdatum oder zur konkreten Naturgefahr erfolgt die Vergabe der Komposit-ID analog zur Vergabe der Ereignis-ID, d. h. jeder Eintrag bekommt eine separate Komposit-ID. Mit anderen Worten werden die Einträge nicht weiter zusammengefasst. Jede CESARE-Ereignis-ID entspricht in diesem Fall einer separaten CESARE-Komposit-ID.

#### Verbleibende Herausforderungen im Bereich Komposit-ID

Weitestgehend analog zu den verbleibenden Herausforderungen im Bereich der Ereignis-ID.

## 4. QUALITÄT DER DATEN UND DES HARMONISIERUNGSPROZESSES

---

Im Zuge der Beschreibungen der Datenharmonisierung in Kapitel 3 wurden bereits immer wieder Anmerkungen zur Qualität der Originaldaten und des Harmonisierungsprozesses eingebracht. Im vorliegenden Kapitel wird diese Information nochmals systematisch aufbereitet und zusammengefasst. Dies betrifft zum einen die Qualität bzw. Güte der Originaldaten und des Harmonisierungsprozesses und zum anderen die auch nach dem Harmonisierungsprozess verbleibenden Herausforderungen.

### QUALITÄT BZW. GÜTE DER ORIGINALDATEN UND DES HARMONISIERUNGSPROZESSES

Grundsätzlich könnte die in De Groeve et al. (2014) beschriebene Pedigree-Methode ein geeignetes Mittel sein, um die Gesamtqualität des Harmonisierungs- und Transformierungsprozesses zu beschreiben. Sie unterzieht (i) das Sammeln und Erfassen der Daten, (ii) das Sortieren und Bearbeiten der Daten und schließlich (iii) das Transformieren der Daten zur Zielerreichung einer Qualitätsbewertung. Das Bewertungsschema ist jedoch sehr detailliert und erfordert beispielsweise Informationen über die Genauigkeit während der Erhebungsphase (Methode, Analyse menschlicher Fehler), über Vergleiche mit unabhängigen Messungen, über die Zuverlässigkeit der Quelle (auf Messungen oder Annahmen beruhende Daten, Vollständigkeit etc.). Eine derartige Bewertung der vorliegenden Datensätze ist mit den verfügbaren Informationen oftmals nicht in dem geforderten Detailgrad möglich. Die Bewertung erfolgt daher in einer etwas abgewandelten und abgespeckten Form.

Tabelle 4-1 gibt einen Überblick über die Qualität bzw. Güte der Originaldaten in Bezug auf die Merkmale, die für das CESARE-Zielschema verwendet werden<sup>6</sup>, sowie über den Harmonisierungsprozess. Sie fasst damit die Informationen aus Kapitel 3 nochmals zusammen. Jedes der für das CESARE-Zielschema verwendeten Merkmale wird dabei anhand der folgenden Kategorien bewertet:

---

<sup>6</sup> Siehe auch Abbildung A-3 in Anhang A für einen Überblick über die aus den Originalquellen extrahierten und für den CESARE-Demonstrator harmonisierten Daten.

- **Informationsgehalt/Detailgrad:** Bewertung der einzelnen Original-Datensätze in Bezug auf den Informationsgehalt bzw. den Detailgrad des jeweils betrachteten Merkmals anhand einer fünfstufigen Skala: niedrig, mittelniedrig, mittel, mittelhoch, hoch. Als „hoch“ wird dabei jener Informationsgehalt bzw. Detailgrad definiert, den die Autor:innen als Optimalfall oder anhand der vorliegenden Datensätze als Best Practice erachten. Die Bewertung der einzelnen Datensätze erfolgt im Vergleich zu diesem Optimal- bzw. Best-Practice-Fall. Ist beispielsweise das Ausmaß an fehlenden Werten („missing values“) erheblich, wird das als Schmälerung des Informationsgehalts erachtet.
- **Unsicherheiten/Unschärfen:** Gibt an, ob im Original-Datensatz in Bezug auf das betrachtete Merkmal Unsicherheiten oder Unschärfen zum Beispiel in Form von Messungenauigkeiten oder variierenden Detailgraden zu verzeichnen sind. Die Bewertung erfolgt anhand der Kategorien „ja“, „nein“, „teilweise“ und „n. b.“ (nicht beurteilbar).
- **Angaben zur Messgenauigkeit:** Gibt an, ob im Original-Datensatz in Bezug auf das betrachtete Merkmal explizite Angaben zur Messgenauigkeit enthalten sind, zum Beispiel in Form der MAXO-Codierung. Die Bewertung erfolgt anhand der Kategorien „ja“, „nein“, „teilweise“ und „n. b.“ (nicht beurteilbar).
- **Harmonisierungspotential:** Bewertung des Harmonisierungspotentials des Original-Merkmals in Bezug auf die Zielvariable, d. h. wie gut lässt sich das Quellschema in das Zielschema transformieren. Die Bewertung erfolgt anhand einer fünfstufigen Skala: niedrig, mittelniedrig, mittel, mittelhoch, hoch. Als „hoch“ wird das Harmonisierungspotential dann eingestuft, wenn sich die jeweils zugrundeliegenden Konzepte gut ineinanderfügen und die Attribute des Quellschemas in eindeutiger Weise den Attributen des Zielschemas zuordnen lassen.
- **Informationsverlust durch Harmonisierung:** Gibt an, ob es durch die Harmonisierung bzw. die Transformation vom Quell- zum Zielschema zu einem Informationsverlust kommt. Die Bewertung erfolgt anhand der Kategorien „ja“, „nein“ und „kaum“. Etwaige Informationsverluste müssen das Harmonisierungspotential nicht unbedingt beeinträchtigen. Verlangt das Zielschema weniger Informationen als im Originalschema vorhanden, kann das Harmonisierungspotential trotz vorhandenen Informationsverlustes hoch sein.

Tabelle 4-1: Bewertung der Qualität und Güte der Originaldaten in Bezug auf die Zielschemata sowie des Harmonisierungsprozesses

	GEORIOS	WLK	HWFDB	VIOLA	Feuerwehr	KatFonds NÖ Privat	KatFonds NÖ Gemeinden	KatFonds Stmk Privat	KatFonds Stmk Land <sup>a)</sup>
<b>Datums-/Zeitangabe</b>									
... Informationsgehalt bzw. Detailgrad	hoch	hoch	hoch	hoch	mittelhoch	niedrig	mittelhoch	mittelhoch	mittelhoch
... Unsicherheiten bzw. Unschärfen	ja	ja	ja	ja	ja	n. b.	ja	n. b.	ja
... Angaben zur Messgenauigkeit	ja	ja	ja	ja	nein	nein	nein	nein	nein
... Harmonisierungspotential	hoch	hoch	hoch	hoch	hoch	hoch	hoch	hoch	hoch
... Informationsverlust durch Harmonisierung	ja	ja	ja	ja	ja	nein	nein	nein	nein
<b>Verortung</b>									
... Informationsgehalt bzw. Detailgrad	hoch	hoch	hoch	mittel	mittelhoch	mittel	mittel	mittel	mittel
... Unsicherheiten bzw. Unschärfen	ja	ja	ja	n. b.	ja	n. b.	n. b.	n. b.	ja
... Angaben zur Messgenauigkeit	nein	nein	nein	nein	ja	nein	nein	nein	nein
... Harmonisierungspotential	hoch	hoch	hoch	hoch	hoch	hoch	hoch	hoch	hoch
... Informationsverlust durch Harmonisierung	ja	ja	ja	nein	ja	nein	nein	ja	ja
<b>Naturgefahr/Hazard</b>									
... Informationsgehalt bzw. Detailgrad	hoch	hoch	mittel	hoch	mittel	mittel	mittel	mittel	mittel
... Unsicherheiten bzw. Unschärfen	n. b.	n. b.	n. b.	n. b.	ja	ja	ja	ja	ja
... Angaben zur Messgenauigkeit	nein	nein	nein	nein	nein	nein	nein	nein	nein
... Harmonisierungspotential	mittelhoch	mittelhoch	hoch	hoch	mittel	mittelhoch	mittel	mittelhoch	mittel
... Informationsverlust durch Harmonisierung	ja	ja	nein	kaum	ja	kaum	kaum	kaum	kaum
<b>Schadenkategorie/Element</b>									
... Informationsgehalt bzw. Detailgrad	mittelniedrig <sup>b)</sup>	hoch	niedrig	mittelhoch	niedrig	mittelhoch	mittelhoch	mittel	hoch

	GEORIOS	WLK	HWFDB	VIOLA	Feuerwehr	KatFonds NÖ Privat	KatFonds NÖ Gemeinden	KatFonds Stmk Privat	KatFonds Stmk Land <sup>a)</sup>
... Unsicherheiten bzw. Unschärfen	n. b.	n. b.	n. b.	n. b.	nein	n. b.	ja	n. b.	ja
... Angaben zur Messgenauigkeit	nein	nein	nein	nein	nein	nein	nein	nein	nein
... Harmonisierungspotential	hoch	mittelhoch	hoch	mittelhoch	hoch	mittelhoch	mittelhoch	mittelhoch	mittelhoch
... Informationsverlust durch Harmonisierung	nein	ja	nein	ja	nein	ja	ja	ja	ja
<b>Schadasmusmaß - Personenschaden</b>									
... Informationsgehalt bzw. Detailgrad	mittel	mittelhoch	-	mittelhoch	-	-	-	-	-
... Unsicherheiten bzw. Unschärfen	ja	ja	-	ja	-	-	-	-	-
... Angaben zur Messgenauigkeit	nein	ja	-	nein	-	-	-	-	-
... Harmonisierungspotential	hoch	hoch	-	hoch	-	-	-	-	-
... Informationsverlust durch Harmonisierung	nein	ja	-	ja	-	-	-	-	-
<b>Schadasmusmaß (monetär) - Sachschaden</b>									
... Informationsgehalt bzw. Detailgrad	-	mittelniedrig	mittelniedrig	mittelniedrig	-	mittelhoch	mittelhoch	mittelhoch	mittelhoch
... Unsicherheiten bzw. Unschärfen	-	ja	ja	ja	-	ja	ja	ja	ja
... Angaben zur Messgenauigkeit	-	nein	ja	nein	-	teilweise	ja	nein	teilweise
... Harmonisierungspotential	-	hoch <sup>c)</sup>	hoch <sup>c)</sup>	- <sup>d)</sup>	-	mittelniedrig	mittelhoch	mittel	mittelhoch
... Informationsverlust durch Harmonisierung	-	Ja	ja	-	-	nein	nein	nein	ja
<b>Schadasmusmaß (monetär) - Einsatzkosten</b>									
... Informationsgehalt bzw. Detailgrad	-	-	-	-	-	-	-	-	mittelhoch
... Unsicherheiten bzw. Unschärfen	-	-	-	-	-	-	-	-	ja
... Angaben zur Messgenauigkeit	-	-	-	-	-	-	-	-	teilweise
... Harmonisierungspotential	-	-	-	-	-	-	-	-	mittelhoch
... Informationsverlust durch Harmonisierung	-	-	-	-	-	-	-	-	ja

	GEORIOS	WLK	HWFDB	VIOLA	Feuerwehr	KatFonds NÖ Privat	KatFonds NÖ Gemeinden	KatFonds Stmk Privat	KatFonds Stmk Land <sup>a)</sup>
<b>Objektbesitzer:in (Sachschaden)</b>									
... Informationsgehalt bzw. Detailgrad	-	mittel	mittel	-	-	hoch	hoch	hoch	hoch
... Unsicherheiten bzw. Unschärfen	-	n. b.	n. b.	-	-	nein	nein	nein	ja
... Angaben zur Messgenauigkeit	-	nein	nein	-	-	-	-	-	nein
... Harmonisierungspotential	-	hoch	hoch	-	-	hoch	hoch	hoch	hoch
... Informationsverlust durch Harmonisierung	-	nein	nein	-	-	nein	nein	nein	nein
<b>Schadenträger:in</b>									
... Informationsgehalt bzw. Detailgrad	-	hoch	hoch	-	-	hoch	hoch	hoch	hoch
... Unsicherheiten bzw. Unschärfen	-	n. b.	n. b.	-	-	ja	n. b.	n. b.	n. b.
... Angaben zur Messgenauigkeit	-	nein	nein	-	-	nein	nein	nein	nein
... Harmonisierungspotential	-	hoch	hoch	-	-	hoch	hoch	hoch	hoch
... Informationsverlust durch Harmonisierung	-	nein	nein	-	-	nein	nein	nein	nein

n. b. = nicht beurteilbar;

<sup>a)</sup> zusätzliche Unsicherheiten durch die Digitalisierung der originalen pdf-Schans, die z. T. händisch zu erfolgen hatte;

<sup>b)</sup> „mittelniedrig“, da die Information in Form von Freitext vorliegt, der im Zuge von CESARE nur z. T. ausgewertet wurde;

<sup>c)</sup> aufgrund des Risikos von Mehrfachnennungen wird nur ein Teil der im Original vorhandenen Informationen zum Schadausmaß übernommen; die Bewertung zum Harmonisierungspotential bezieht sich auf den Teil der Information, der übernommen wird;

<sup>d)</sup> wird nicht in das CESARE-Zielschema übernommen aufgrund des Risikos von Mehrfachnennungen

## VERBLEIBENDE HERAUSFORDERUNGEN TROTZ HARMONISIERUNG

In CESARE werden unterschiedliche Datenquellen verwendet, die Ereignis- und Schadendaten aus unterschiedlichen Zwecken, nach unterschiedlichen Ansätzen und in unterschiedlichem Umfang sammeln. Daraus ergeben sich viele Herausforderungen für die Schaffung eines möglichst homogenen Datensatzes. Die meisten davon überschneiden sich mit den von Faiella et al. (2020) identifizierten Herausforderungen, die mit dem Risk Data Hub ein analoges Unterfangen auf europäischer Ebene vorantreiben. Zu den in CESARE identifizierten Herausforderungen zählen unter anderem (siehe auch Abbildung 4-1):

- Unterschiede im Zeitpunkt des Auftretens bzw. Unschärfen im Ereignis-/Schadenzeitpunkt
- Unterschiede in der Klassifizierung der räumlichen Ausdehnung und der Detailliertheit der Verortung sowie Unschärfen in der Verortung
- Unterschiede und Unschärfen in der Definition der Attribute bzw. dem verwendeten Vokabular
- Unterschiede in der Ereignisabgrenzung bzw. -definition (Art der Naturgefahr)
- Unterschiede in den Definitionen der Indikatoren, z. B. unterschiedliche Kriterien, ab wann ein Schaden berücksichtigt bzw. dokumentiert wird
- Unterschiede in den Bewertungsmethoden von Schäden

Abbildung 4-1: Herausforderungen in der Harmonisierung und Standardisierung



Quelle: eigene Darstellung; Icons von Freepik ([www.flaticon.com](http://www.flaticon.com))

Viele dieser Herausforderungen führen zu einem gewissen Risiko von Mehrfachzählungen beim Zusammenführen und/oder schränken die Vergleichbarkeit und Harmonisierungsmöglichkeit ein. Ein Teil dieser Herausforderungen lässt sich im Zuge des Harmonisierungsprozesses abschwächen, jedoch nicht alle. Somit finden sich viele

Unschärfen der Original-Datensätze auch im harmonisierten Datensatz wieder. Folgende Herausforderungen und Unschärfen verbleiben auch nach dem Harmonisierungsprozess:

- Unterschiede in der Ereignisabgrenzung und Unschärfen in Bezug auf die Datums-/Zeitangabe, Verortung sowie Naturgefahr führen dazu, dass auch nach erfolgter Zusammenführung Ereignisse mehrfach im harmonisierten Datensatz aufscheinen können. Eine Auswertung des zusammengeführten CESARE-Datensatzes in Bezug auf die Anzahl der Ereignisse ist daher nur bedingt sinnvoll und aussagekräftig.
- In den Original-Datensätzen vorhandenen Unschärfen in Bezug auf die Abgrenzung von Naturgefahren wurden in das Zielschema mitübernommen. Dasselbe gilt für Unschärfen in der Abgrenzung von Schadenkategorien (geschädigtes Element).
- Vor allem die KatFonds-Datensätze können je nach Bundesland unterschiedliche Regelungen aufweisen, welche Schäden beihilfenfähig sind und daher dokumentiert werden. Diese Unterschiede bleiben auch im harmonisierten CESARE-Datensatz bestehen.

Die verbleibenden Herausforderungen und Unsicherheiten sind bei der Nutzung des harmonisierten CESARE-Datensatzes mitzudenken. Ebenfalls mitzudenken sind der Umfang und die Vollständigkeit der bislang erfassten Schäden. Hauptinformationsquelle für monetäre Schäden stellen derzeit die KatFonds-Datensätze dar, angereichert durch Informationen zu Schäden an Verbauungs- und Schutzinfrastruktur der Wildbach- und Lawinerverbauung (WLV) und Bundeswasserbauverwaltung (BWV) aus dem WLK- und HWFDB-Datensatz. Damit ist bislang nur ein Teil der insgesamt anfallenden Schäden erfasst. Zu den nicht berücksichtigten Schäden zählen u. a.:

- **Schäden an Betrieben unterhalb der Existenzgefährdung:** Voraussetzung für Beihilfen durch Länder und Bund (Katastrophenfonds) ist eine spürbare materielle Belastung. Im Falle juristischer Personen muss eine Existenzgefährdung vorliegen. Große Unternehmen sind daher selten vom Katastrophenfonds und damit zusammenhängenden Schadendatensammlungen (KatFonds-Datensätze) erfasst.
- **Versicherte Schäden:** Ein wesentlicher Teil, der für einen guten Überblick über das gesamte SchADVolumen derzeit fehlt, sind versicherte Schäden. Diese sind i. d. R nicht in den KatFonds-Datensätzen enthalten bzw. aus den dort angeführten anerkannten Schäden herausgerechnet, da versicherte Schäden nicht beihilfenfähig sind. Je nach Naturgefahr decken die KatFonds-Datensätze daher auch einen unterschiedlich hohen Anteil des Gesamtschadenvolumens ab. So sind Sturmschäden an und im Haus üblicherweise über die Eigenheimversicherung und die Haushaltsversicherung gedeckt. Durch die hohe Marktdurchdringungsrate von Eigenheim- und Haushaltsversicherung dürften daher die bei den Versicherern vorhandenen Informationen über weite Teile die Sturmschäden an und in Häusern abdecken, während die KatFonds-Datensätze hierzu sehr lückenhaft und nicht repräsentativ sind. Letztere machen im Gebäudebereich schätzungsweise weniger als 10 % der versicherten Sturmschäden aus.

Die normale Sturmversicherung für Eigenheime deckt neben Sturm- auch Hagelschäden sowie Schäden durch Erdbeben und Felssturz. Nicht im Rahmen der normalen Sturmversicherung gedeckt sind hingegen Schäden durch Überschwemmung, Hochwasser oder Vermurung. Für diese Gefahren bedarf es einer gesonderten Katastrophenabdeckung, die nur eingeschränkte Versicherungssummen enthält. Gebäudeschäden durch Hochwasser und Vermurung werden von den Katastrophenfondsdaten also deutlich umfassender abgedeckt als etwa Gebäudeschäden durch Sturm.

- **Schäden am Vermögen des Bundes:** Abgesehen von Schäden an der Verbauungs- und Schutzinfrastruktur der WLW und BWV sind derzeit keine Schäden am Vermögen des Bundes berücksichtigt.

Auch wenn die im Rahmen von CESARE verfügbaren Datensätze noch bei weitem nicht das gesamte SchADVolumen für die drei hauptsächlich betrachteten Naturgefahren (Hochwasser, Sturm, Massenbewegungen) erfassen, konnten mit ihrer Hilfe Konzepte erstellt und erprobt werden, wie man Daten aus unterschiedlichen Quellen transformieren und zusammenführen kann, sodass eine gemeinsame Auswertung – unter Berücksichtigung verbleibender Unsicherheiten und Unschärfen – möglich ist. Ein Beispiel für den Informationsgewinn, den eine solche Zusammenführung bringen kann, ist in Anhang A (Abbildung A-4 bis Abbildung A-9) dargestellt. Für ein kombiniertes Ereignis zwischen 11. und 15. 09. 2014 in der Steiermark, bestehend aus durch Starkregen ausgelöste Massenbewegungen und Hochwasserereignisse, illustriert das Beispiel den je Datensatz hinzukommenden Informationsgewinn und damit den Mehrwert, der durch die gemeinsame Nutzung der Daten entstehen kann.

## 5. REFERENZEN

---

- Amt der Niederösterreichischen Landesregierung, 2005. Richtlinien für die Gewährung von Beihilfen zur Behebung von Katastrophenschäden (novelliert 2016).
- Amt der Steiermärkischen Landesregierung, 2018. Richtlinie für die Abwicklung des Entschädigungsverfahrens nach Katastrophenschäden im Vermögen natürlicher und juristischer Personen mit Ausnahme der Gebietskörperschaften im Bundesland Steiermark – Katastrophenfonds-Richtlinie Steiermark.
- Amt der Steiermärkischen Landesregierung, 2013. Richtlinie über die Abwicklung von Soforthilfe- und Folgemaßnahmen der Katastrophenschutzbehörde im Katastrophenfall.
- De Groeve, T., Poljansek, K., Ehrlich, D., Corbane, C., 2014. Current Status and Best Practices for Disaster Loss Data recording in EU Member States: A comprehensive overview of current practice in the EU Member States. Joint Research Centre, Luxembourg.
- Faiella, A., Antofie, T.-E., Luoni, T., Rios Diaz, F., Marin Ferrer, M., 2020. The Risk Data Hub loss datasets - The Risk Data Hub Historical Event Catalogue, JRC Technical Report. Luxembourg.
- JRC, 2015. Guidance for Recording and Sharing Disaster Damage and Loss Data - Towards the development of operational indicators to translate the Sendai Framework into action, JRC Science and Policy Reports. Joint Research Centre, Luxembourg.
- Köberl, J., Prettenhaler, F., Schubert, C., 2018. DAMAGE.at - Machbarkeitsanalyse des Aufbaus einer österreichweiten Schadendatenbank zu wetter- und klimabedingten Infrastrukturschäden, Endbericht im Auftrag des Klima- und Energiefonds. Graz.
- MRC, 2017. DAPA Measurement Toolkit [WWW Document]. URL <https://dapa-toolkit.mrc.ac.uk/concepts/harmonisation> (accessed 3.9.21).
- Statistik Austria, 2020. Gemeindeänderungen ab 1945.

## A. ANHANG

---

Abbildung A-1: Entscheidungsbaum zur Ereignisidentifikation bzw. Ereignis-ID-Vergabe

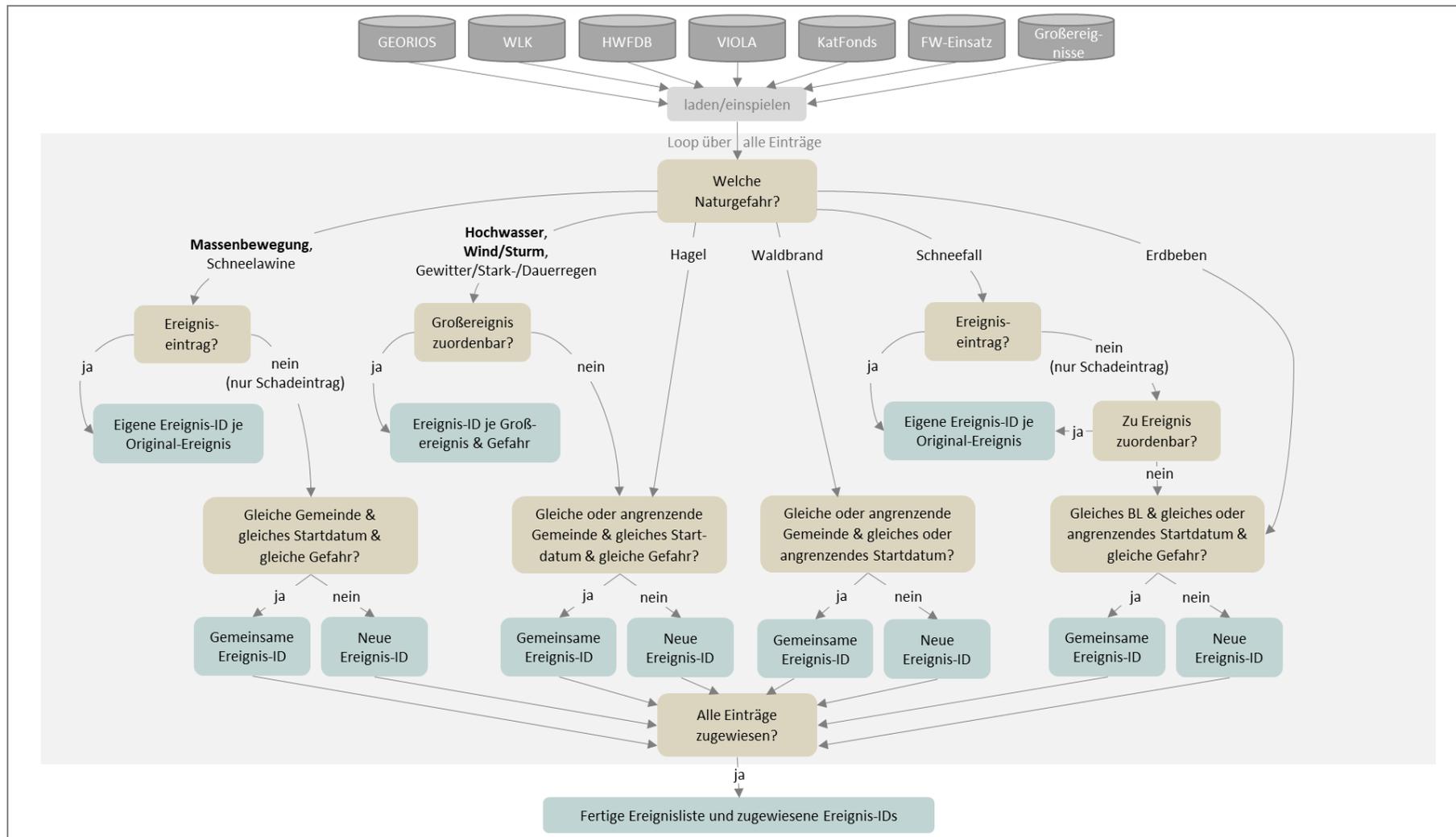


Abbildung A-2: Entscheidungsbaum zur Komposit-ID-Vergabe

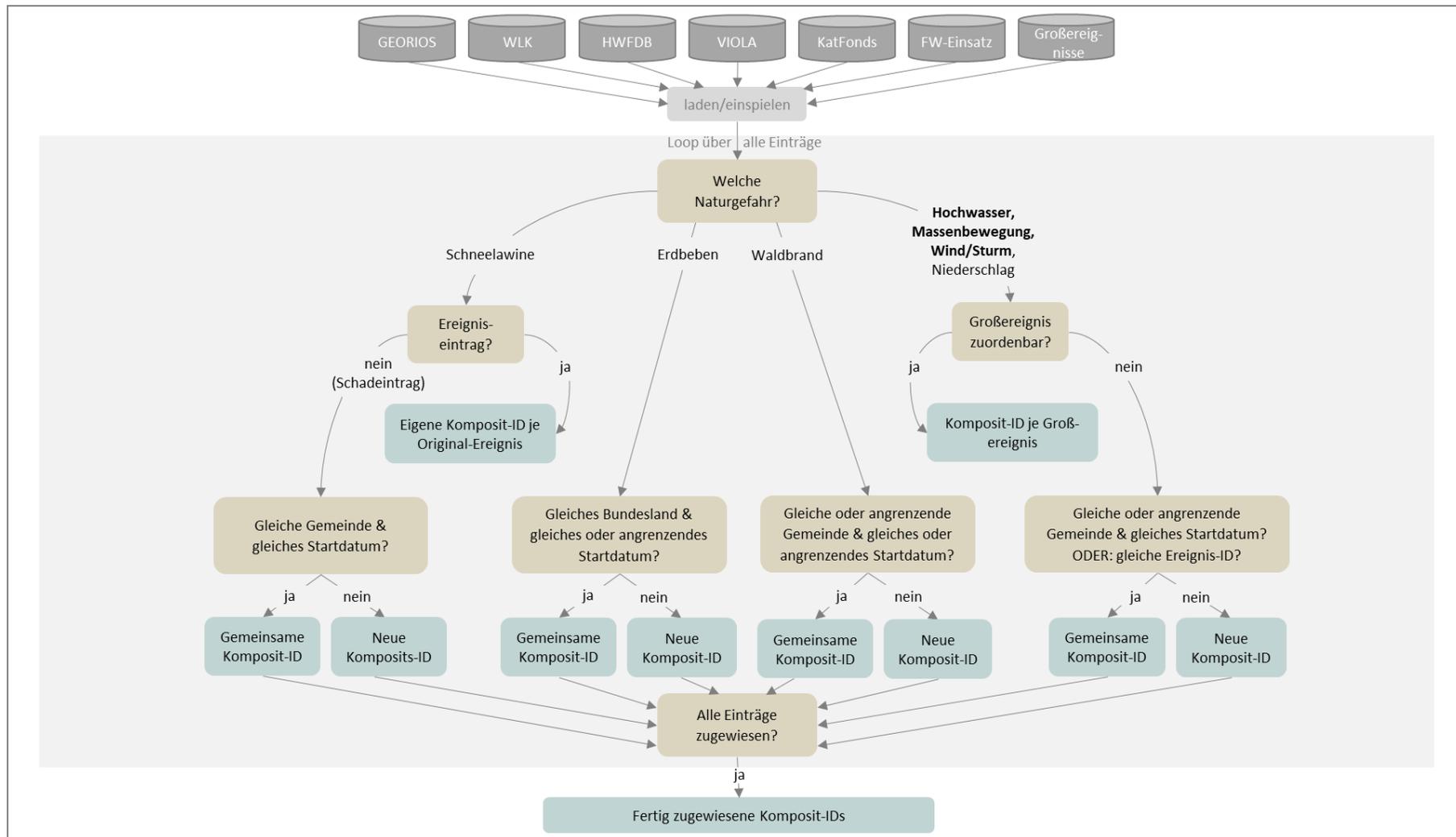


Abbildung A-3: Aus Originalquellen extrahierte und für den CESARE-Demonstrator harmonisierte Daten

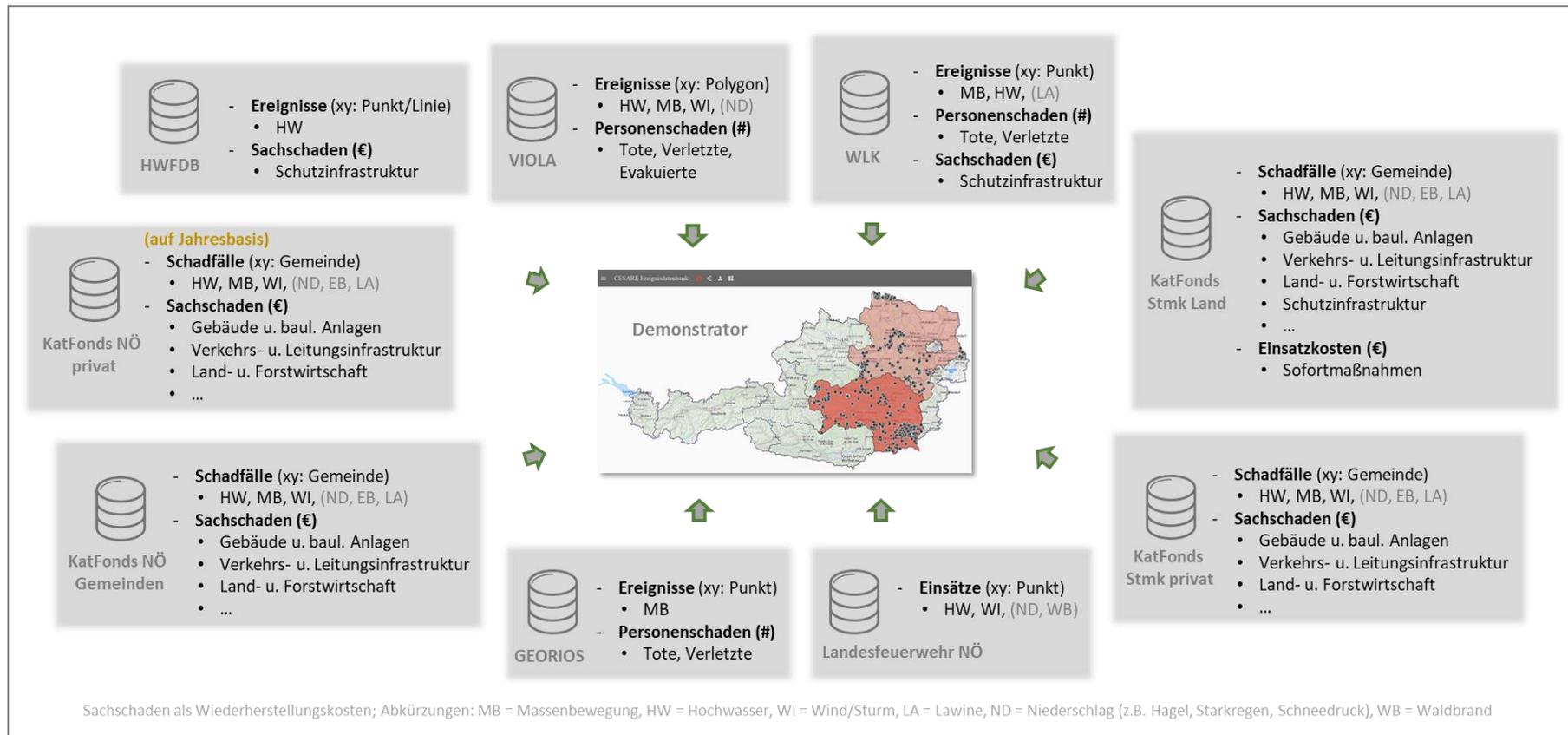


Abbildung A-4: Informationsgewinn durch zusätzliche Datenquellen – Daten vom Land Steiermark zu beihilfenfähigen Schäden im Privatvermögen

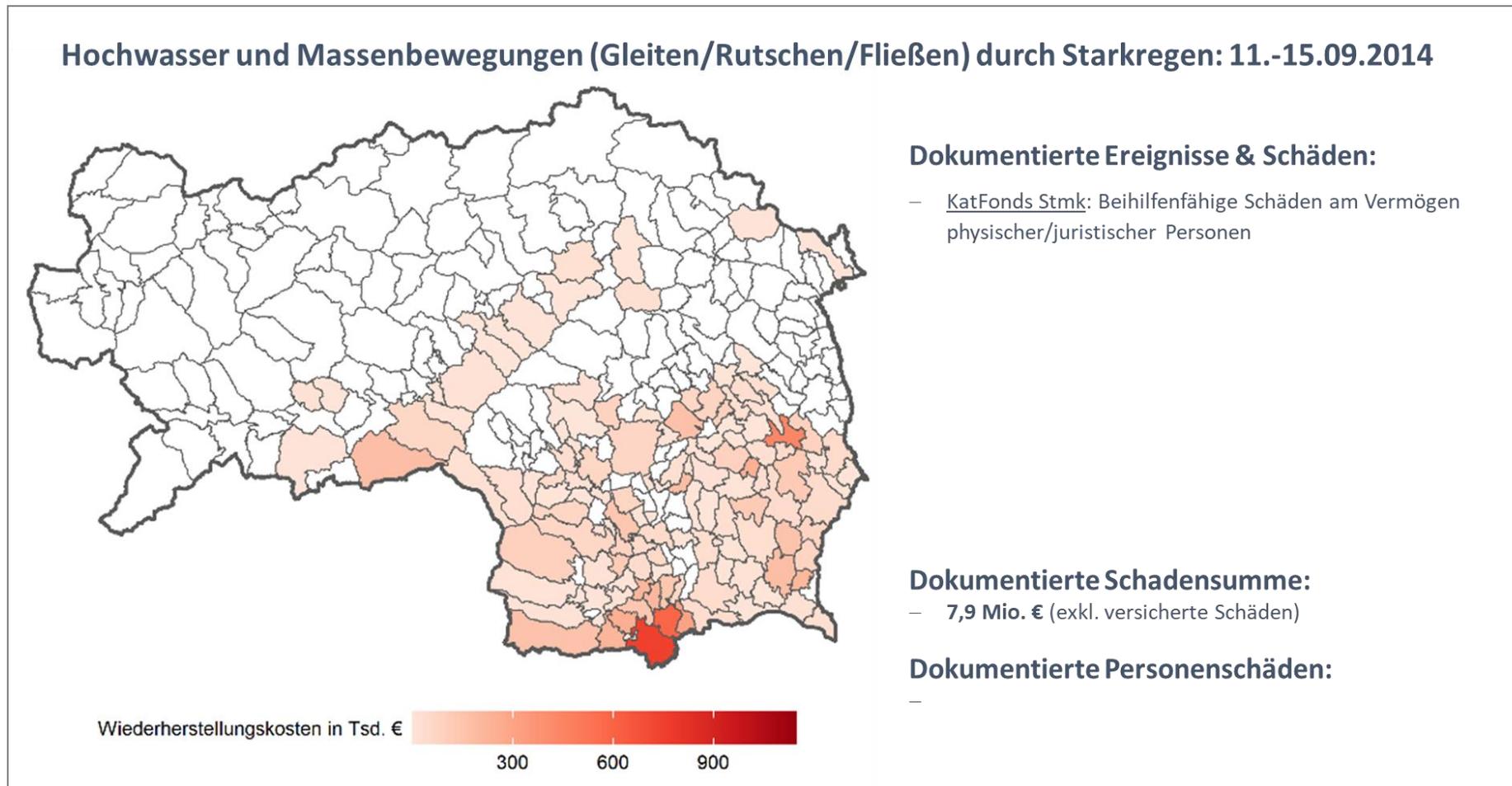


Abbildung A-5: Informationsgewinn durch zusätzliche Datenquellen – PLUS Daten vom Land Steiermark zu Schäden im Landesvermögen

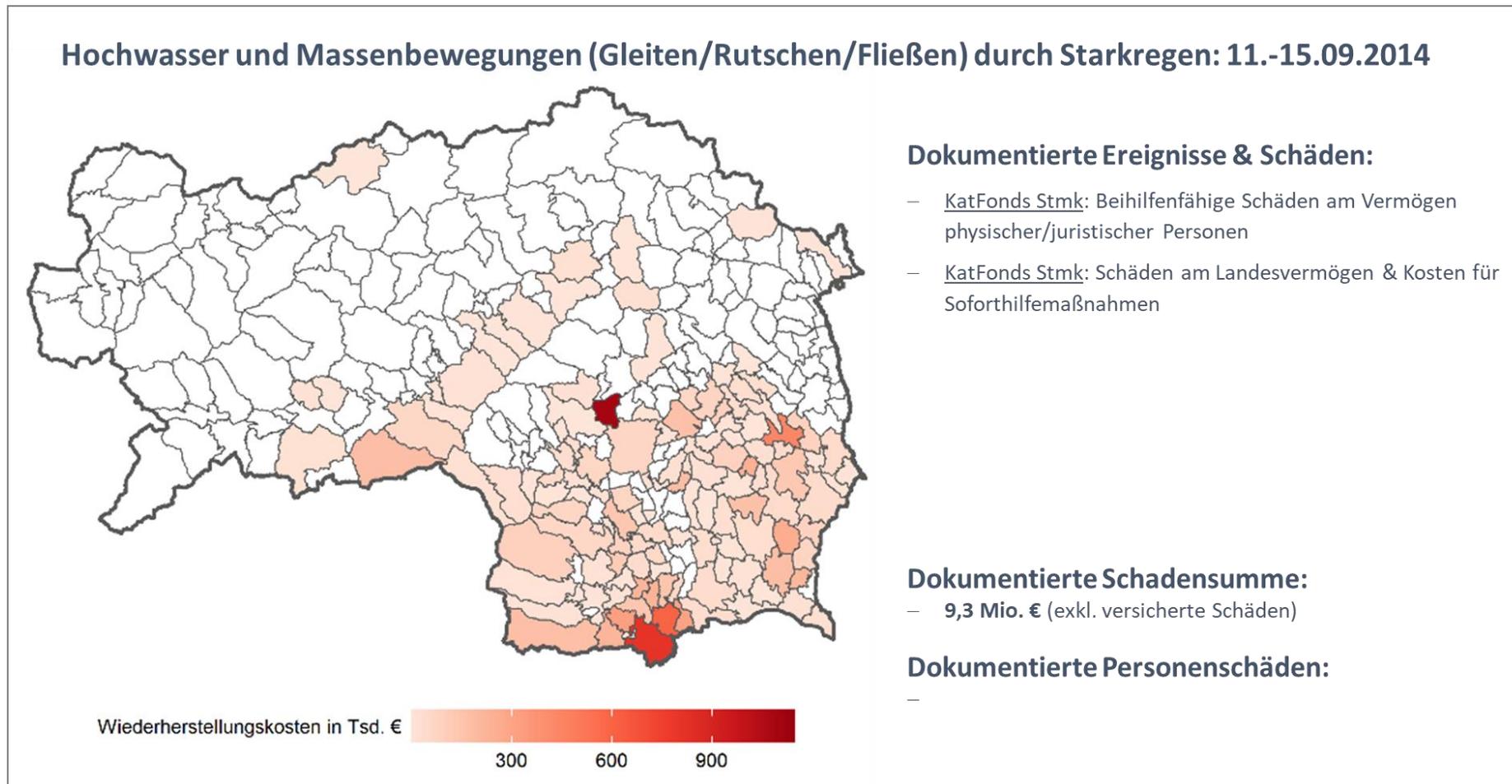


Abbildung A-6: Informationsgewinn durch zusätzliche Datenquellen – PLUS Daten aus der HWFDB zu Hochwasserereignissen und Schäden an Schutzinfrastruktur

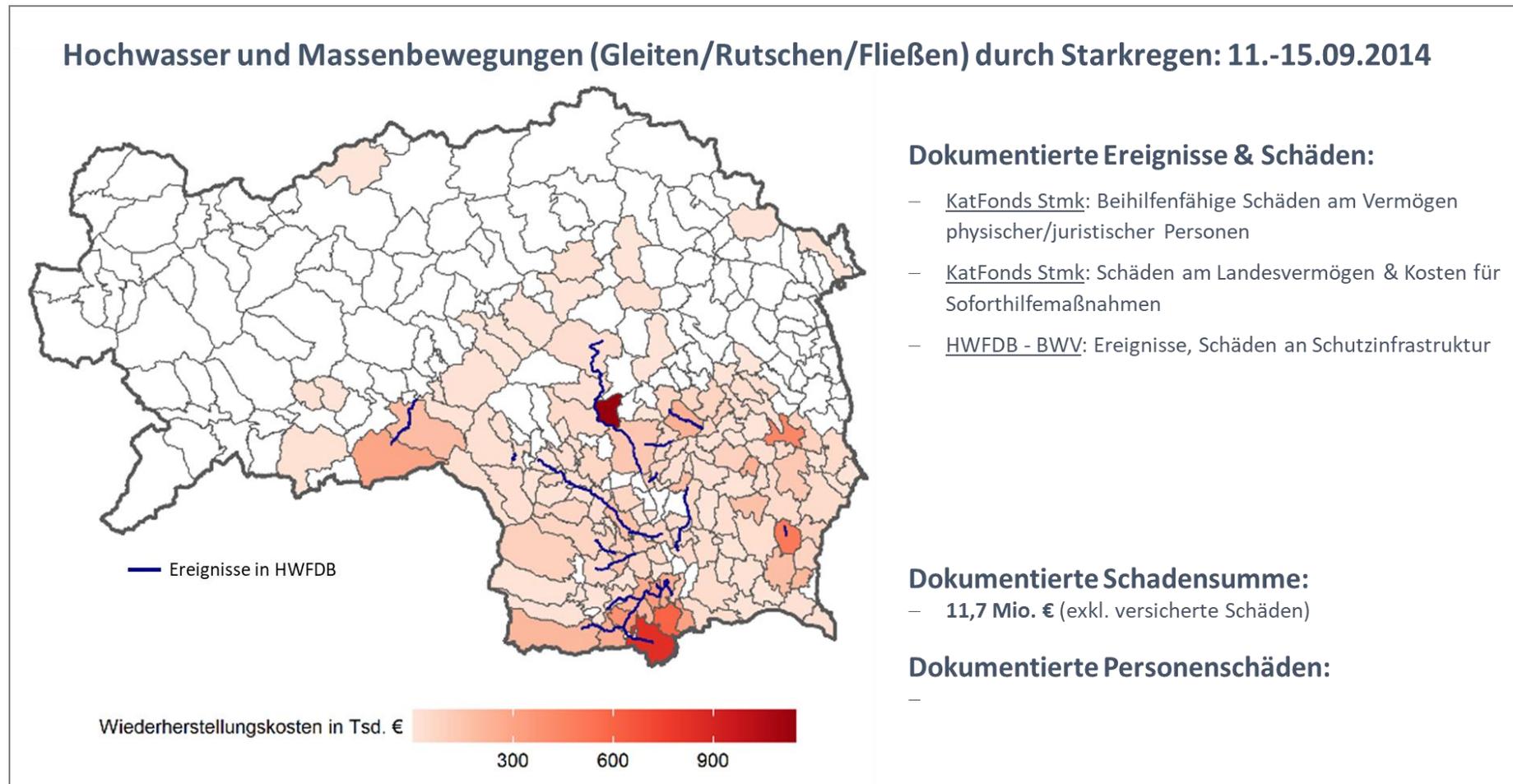


Abbildung A-7: Informationsgewinn durch zusätzliche Datenquellen – PLUS Daten aus dem WLK zu Wildbachereignissen und Schäden an Schutzinfrastruktur

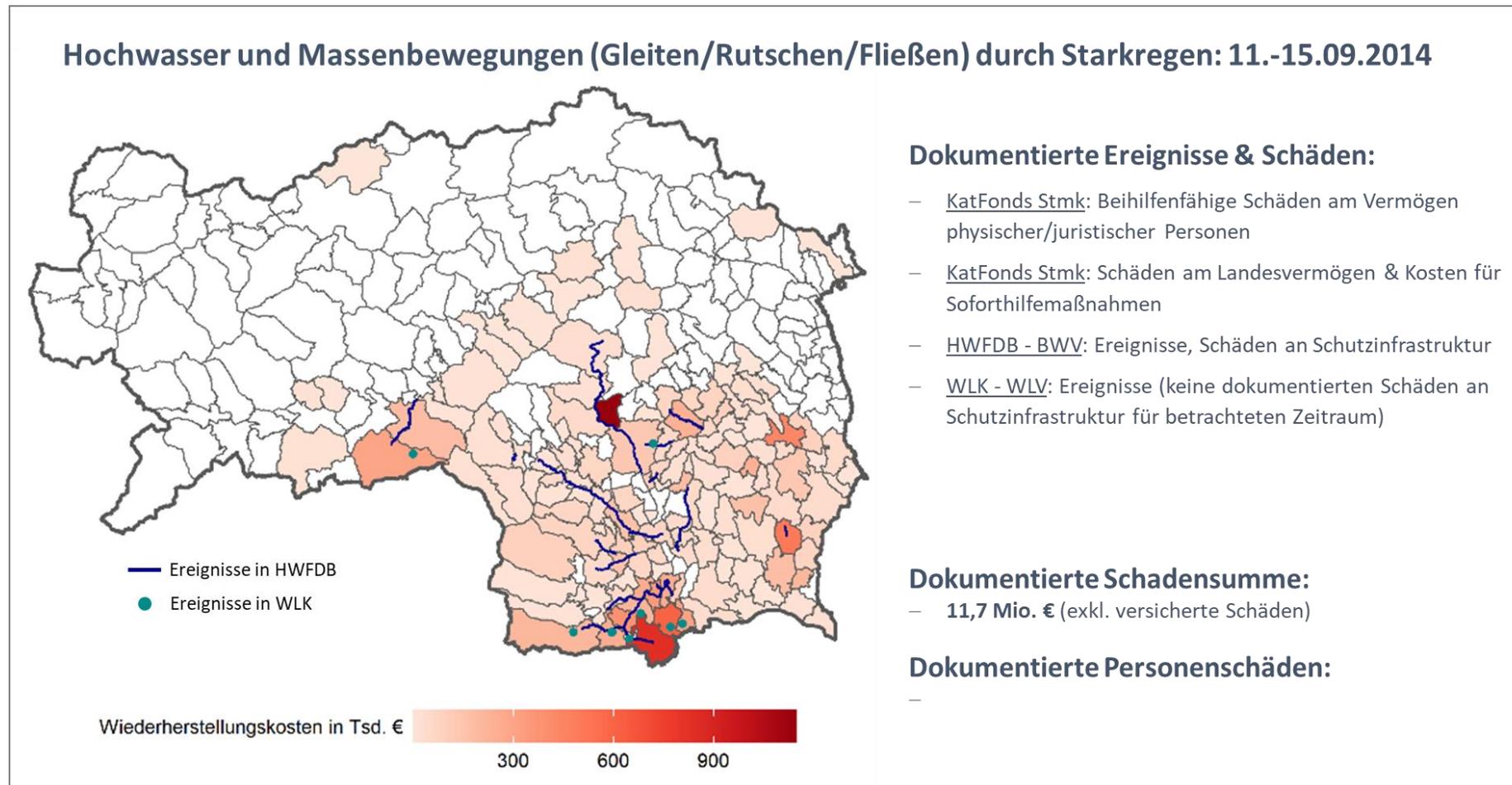


Abbildung A-8: Informationsgewinn durch zusätzliche Datenquellen – PLUS Daten aus GEORIOS zu Massenbewegungsereignissen

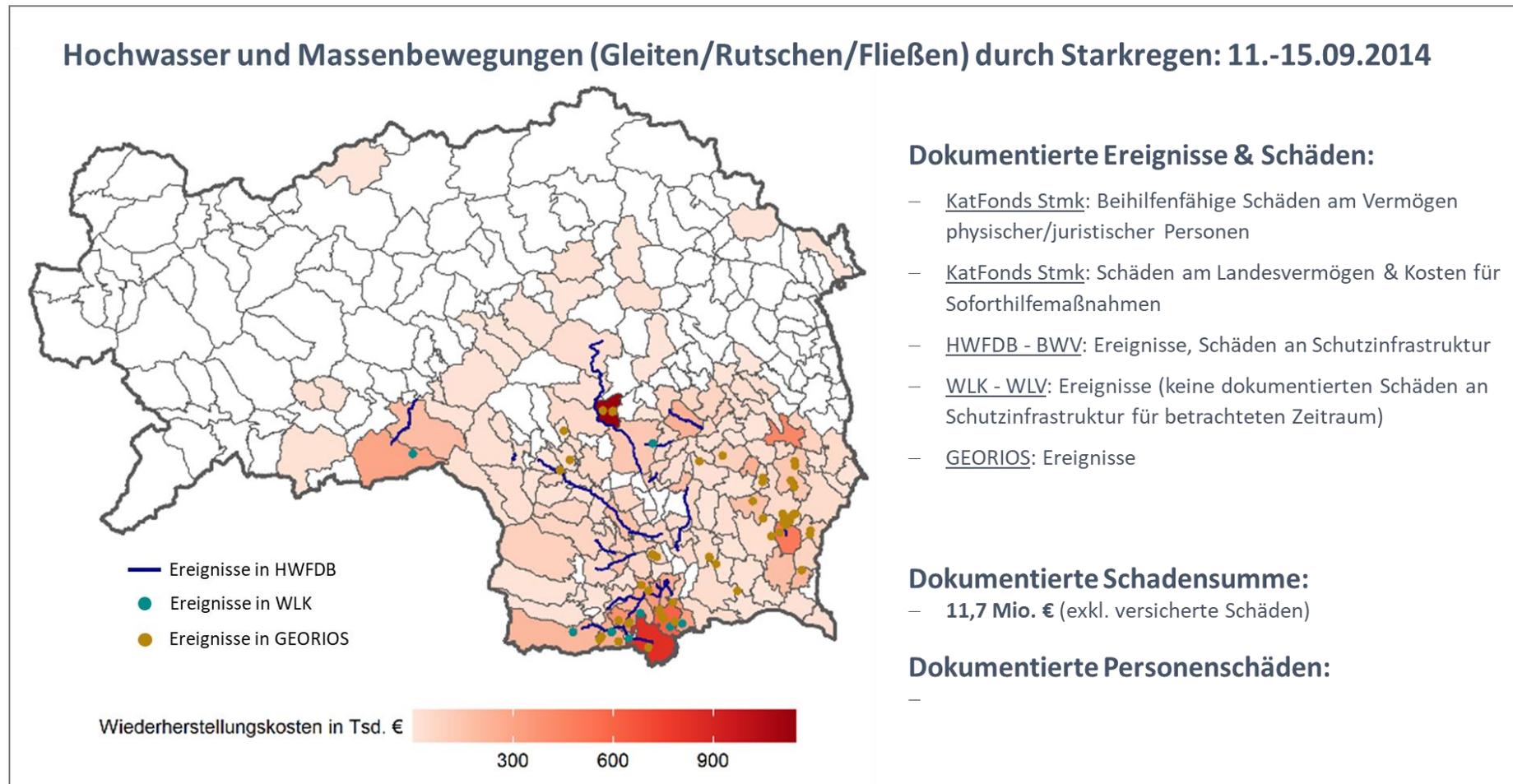


Abbildung A-9: Informationsgewinn durch zusätzliche Datenquellen – PLUS Daten aus VIOLA zu Personenschäden

