Shibboleth ohne Tracking bei den Verlagen

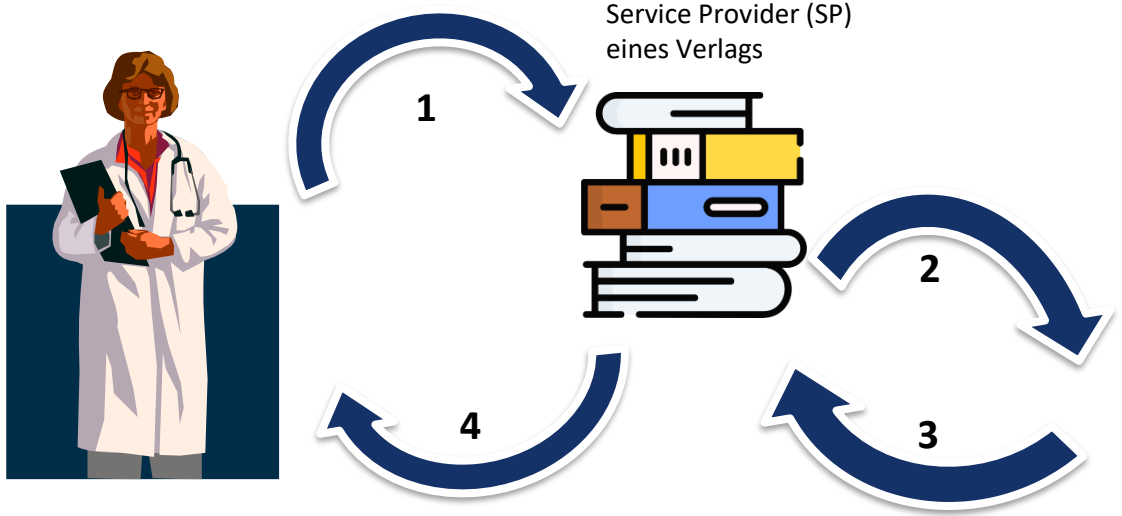
Maria Huber und Julika Mimkes (Gruppe Digitale Medien, SUB Göttingen, Deutschland)

Rahmenbedingungen an der Uni Göttingen

- Die SUB Göttingen gewährt bisher Zugriff auf elektronische Medien von außen über HAN und VPN.
- Mit der Einführung eines Identity Managements bei der GWDG (dem Rechenzentrum u.a. für die Uni Göttingen) sind die lizenzrechtlichen Rahmenbedingungen gegeben, um auch „Shibboleth“ einzuführen.
- Anstelle von Shibboleth wird die ebenfalls SAML2-standardkonforme Software **SimpleSAMLphp** (https://simplesamlphp.org/) für den Göttinger Identiy Provider verwendet. Im Vortrag wird deshalb die Schreibweise „Shibboleth“ genutzt.
- In 2023 wird „Shibboleth“ für den Zugriff auf Verlagspublikationen eingeführt.

Ziel: „Shibboleth“ ohne Tracking im Verlagskontext

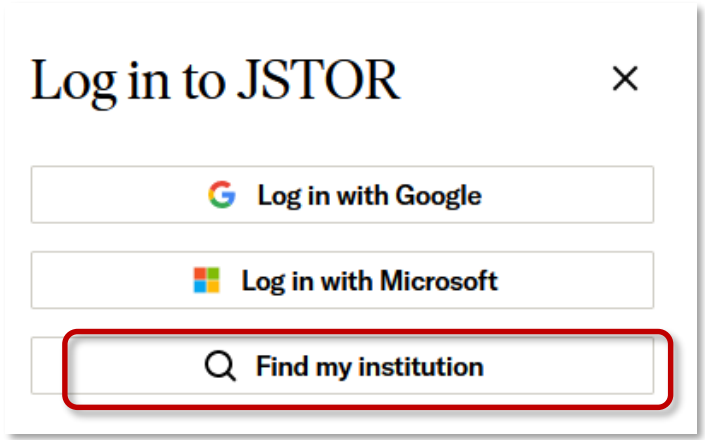
- „Shibboleth“ soll als Single Sign-on-Mechanismus im Verlagskontext eine anonymisierte Authentifizierung ohne Möglichkeit zum verlagsseitigen Tracking anbieten.
- Wir orientieren uns an den Empfehlungen der deutschen Shibboleth-Förderation DFN-AAI (https://www.aai.dfn.de).
- Empfehlungen und Guidelines der FIM4L Gruppe: Principles & Recommendations for Library Services https://www.fim4l.org/?page_id=257 (version 1.0 von 2020, current version 2.0 von 2022)

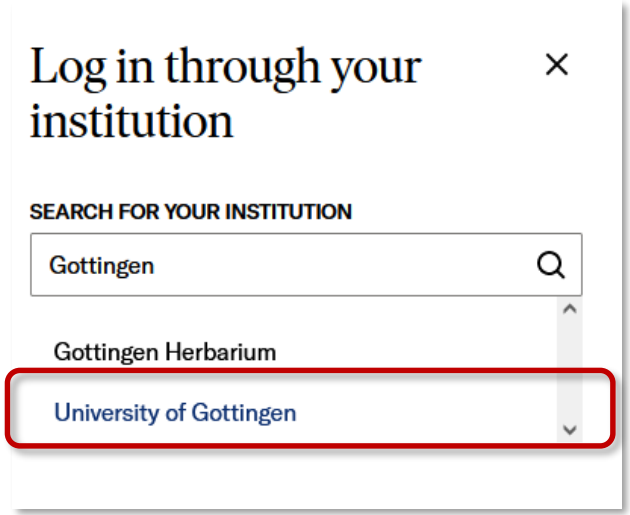


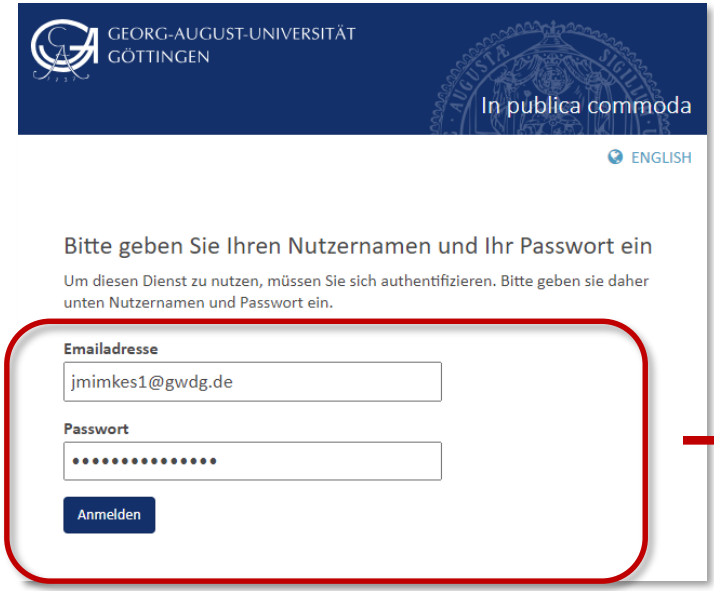
Shibboleth

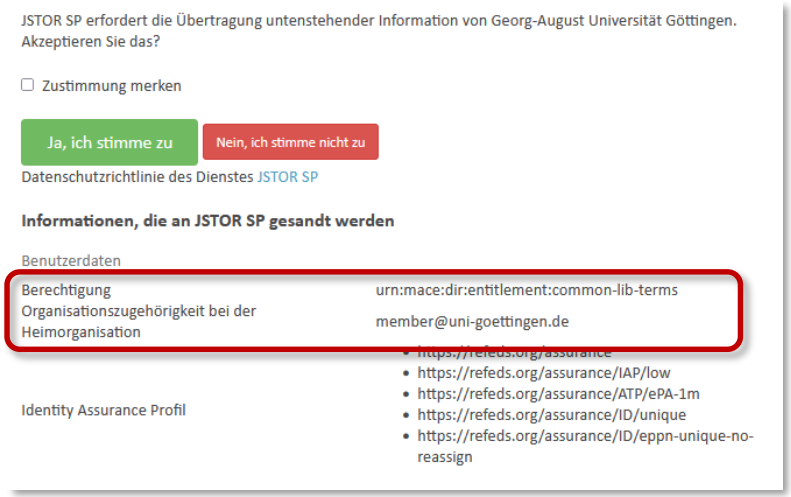
GWDG
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen

1. Eine Nutzerin möchte einen Artikel lesen, dazu begibt sie sich auf die Webseite eines Verlages.
2. Der Verlag bietet ein „Institutionelles Login“ an. Dazu muss sich die Nutzerin beim Identity Provider der Uni Göttingen (gehostet von der GWDG) **mit ihrem Uni-Account authentifizieren**.
3. Der Identity Provider bestätigt dem Verlag, dass die Nutzerin zur Universität Göttingen gehört. **Die übertragenen Informationen vom IdP an den SP sollten nicht personenbezogen sein**.
4. Die Nutzerin wird **auf der Plattform des Verlags autorisiert** und kann Artikel lesen/herunterladen.


Log in to JSTOR
Log in with Google
Log in with Microsoft
Find my institution


Log in through your institution
SEARCH FOR YOUR INSTITUTION
Gottingen
Gottingen Herbarium
University of Gottingen


GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN
In publica commoda
ENGLISH
Bitte geben Sie Ihren Nutzernamen und Ihr Passwort ein
Um diesen Dienst zu nutzen, müssen Sie sich authentifizieren. Bitte geben sie daher unten Nutzernamen und Passwort ein.
Emailadresse
jmimkes1@gwdg.de
Passwort
Anmelden


JSTOR SP erfordert die Übertragung untenstehender Information von Georg-August Universität Göttingen. Akzeptieren Sie das?
Zustimmung merken
Ja, ich stimme zu
Nein, ich stimme nicht zu
Datenschutzrichtlinie des Dienstes JSTOR SP
Informationen, die an JSTOR SP gesandt werden
Benutzerdaten
Berechtigung
urn:mace:dir:entitlement:common-lib-terms
Organisationszugehörigkeit bei der Heimorganisation
member@uni-goettingen.de
Identity Assurance Profil
https://refeds.org/assurance/IAP/low
https://refeds.org/assurance/ATP/ePA-1m
https://refeds.org/assurance/ID/unique
https://refeds.org/assurance/ID/eppn-unique-no-reassign

Konfiguration des IdP der Uni Göttingen

- Nur für Uni-Angehörige (Studierende, Mitarbeitende, Gäste) werden vom IdP Attribute mitgegeben, die bei den Verlagsdiensten für die Autorisierung ausgewertet werden können.
- Bei diesen Attributen handelt es sich um anonyme, datensparsame Informationen:
  - Berechtigung **eduPersonEntitlement** mit Wert **urn:mace:dir:entitlement:common-lib-terms**
  - Zugehörigkeit **eduPersonScopedAffiliation** mit Wert **member@uni-goettingen.de**
- Um den Arbeitsaufwand gering zu halten, werden diese Informationen an alle anfragenden Service Provider / Verlage herausgegeben.
- Für die meisten Anbieter können wir diese datensparsamen Anforderungen recht einfach erfüllen.

Empfehlung der Föderation DFN-AAI zur Attributfreigaben für Verlagsanbieter:
https://doku.tid.dfn.de/doku.php?id=de:shibidp:config-attributes-publishers&rev=1646906103

Code of Conduct

- Der „GÉANT Data Protection Code of Conduct for Service Providers in EU/EEA (CoCo)“ ist eine Selbstverpflichtungserklärung, die den Umgang mit im Shibboleth-Kontext übertragenen personenbezogenen Daten gemäß den geltenden Datenschutzrichtlinien regelt.
- Anbieter, die dem "Code of Conduct" beigetreten sind, erhalten von unserem IdP aktuell zusätzliche personenbezogene Daten, wie den Namen, sofern ihre SP diese anfordern.
- Es sind auch einige Verlagsanbieter Mitglied des "Code of Conduct", was ihnen dadurch Tracking und die Erstellung von Nutzerprofilen ermöglicht.

Informationen zum Code of Conduct:

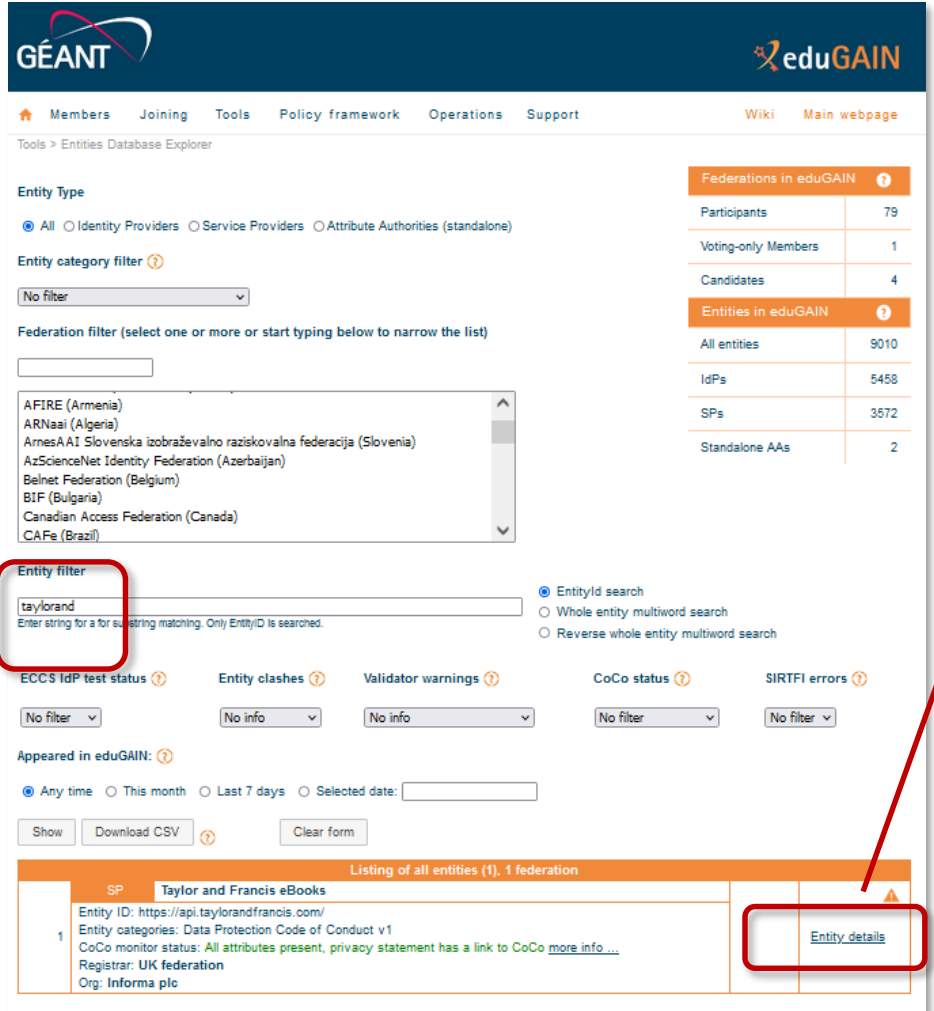
https://doku.tid.dfn.de/doku.php?id=de:geant_coco&rev=1698847675
http://www.geant.net/uri/dataprotection-code-of-conduct/V1 (CoCo Version 1)
https://wiki.refeds.org/display/CODE (CoCo Version 2)

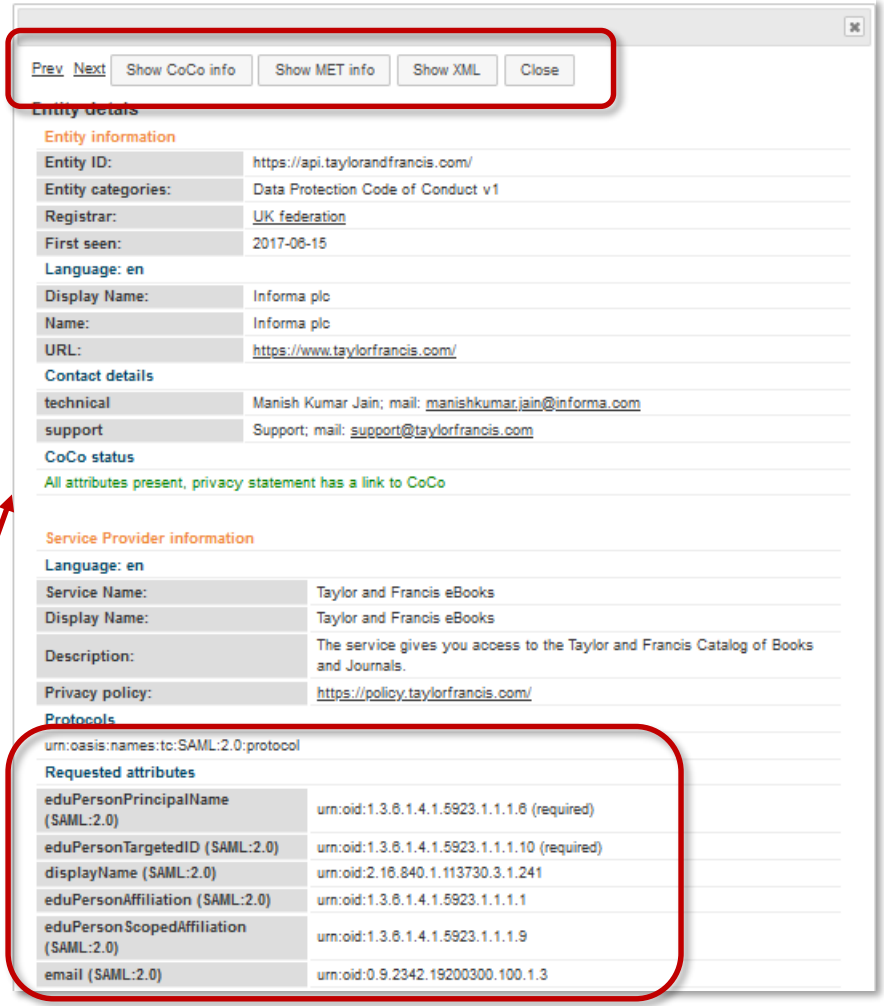
Beispiele für Verlage im Code of Conduct

- *Cambridge University Press / Cambridge CORE*: http://www.cambridge.org/core
  requested attributes: eduPersonScopedAffiliation
- *IOP Publishing / IOP Science*: http://iopscience.iop.org
  requested attributes: eduPersonScopedAffiliation
- *Taylor & Francis eBooks*: https://www.taylorfrancis.com
  requested attributes: **eduPersonPrincipalName, eduPersonTargetedID, displayName,** eduPersonAffiliation, eduPersonScopedAffiliation, **email**
- *Trans Tech / Scientific.Net*: http://www.scientific.net
  requested attributes: eduPersonEntitlement, eduPersonScopedAffiliation

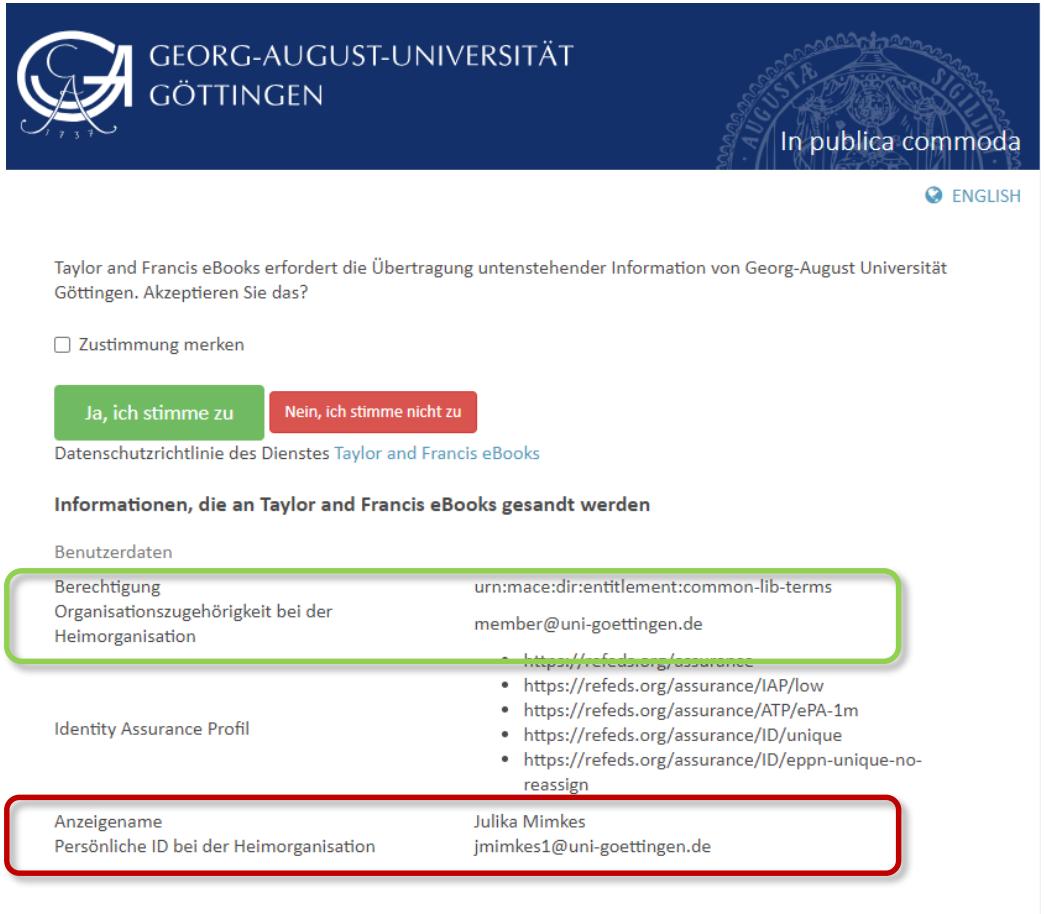
Service Provider mit Entity category http://www.geant.net/uri/dataprotection-code-of-conduct/v1
Quelle: eduGAIN-Monitor https://monitor.edugain.org/coco

Welche Attibute werden von den SP angefragt? Der eduGAIN Entities Database Explorer https://technical.edugain.org/entities





Attributfreigabe beim Login am IdP (Beispiel T&F eBooks)



Browser Add-on SAML Tracer (Beispiel T&F eBooks)

Z. B. mit dem Firefox Add-on **SAML-tracer** können die vom IdP an den SP übermittelten Informationen kontrolliert werden.
(Bei verschlüsselten Übertragungen ist dies allerdings nicht möglich.)

In diesem Beispiel kann man nicht nur die in der Attributfreigabe angezeigten Attribute sehen, sondern auch **versteckt weitergegebene Attribute** sowie weitere Informationen.

SAML 2.0 AttributeStatement

urn:oid:1.3.6.1.4.1.5923.1.1.1.7	urn:mace:dir:entitlement:common-lib-terms
urn:oid:1.3.6.1.4.1.5923.1.1.1.9	member@uni-goettingen.de
urn:oid:1.3.6.1.4.1.5923.1.1.1.11	https://refeds.org/assurance
urn:oid:1.3.6.1.4.1.5923.1.1.1.11	https://refeds.org/assurance/IAP/low
urn:oid:1.3.6.1.4.1.5923.1.1.1.11	https://refeds.org/assurance/ATP/ePA-1m
urn:oid:1.3.6.1.4.1.5923.1.1.1.11	https://refeds.org/assurance/ID/unique
urn:oid:1.3.6.1.4.1.5923.1.1.1.11	https://refeds.org/assurance/ID/eppn-unique-no-reassign
urn:oid:2.16.840.1.113730.3.1.241	Julika Mimkes
urn:oid:1.3.6.1.4.1.5923.1.1.1.6	jmimkes1@uni-goettingen.de
urn:oid:1.3.6.1.4.1.5923.1.1.1.10	740f78ce9c95cb9d3eae0087a0e2ae5f44dedc5c

eduPersonTargetedID: OID 1.3.6.1.4.1.5923.1.1.1.10
Pseudonymer Identifier für eine Person
Ein eindeutiges, dauerhaftes Pseudonym einer Person für einen speziellen Anbieter (pairwise und persistent).

Weiteres Problem: Pairwise persistent ID (Attribute)

- Eine pairwise persistent ID wird IdP-seitig eindeutig pro Nutzerkonto und SP automatisch generiert.
  Sie ist ein eindeutiges, dauerhaftes Pseudonym einer Person für einen speziellen SP, der nicht neu an andere Nutzer*innen vergeben werden darf.
- Mit einer "pairwise persistent“ ID können (zunächst anonyme) Personenprofile beim Verlag erstellt werden, da bei jedem Shibboleth-Login einer Person immer die gleiche ID an den Verlag weitergegeben wird.
- Bereits durch ein einmaliges persönliches Login direkt an der Plattform könnte beim SP/Verlag der eigene Name mit der persistent ID dauerhaft verknüpft werden.
- **Eine pairwise persistent ID ermöglicht Tracking innerhalb eines SP**, aber kein SP-übergreifendes Tracking.
- Beispiele für „pairwise persistent“ Attribute:
  - **eduPersonTargetedID**
  - Die **Pairwise Id** soll die eduPersonTargetedID ablösen.

Informationen zu Attributen (DFN-AAI): https://doku.tid.dfn.de/doku.php?id=de:common_attributes&rev=1687527663

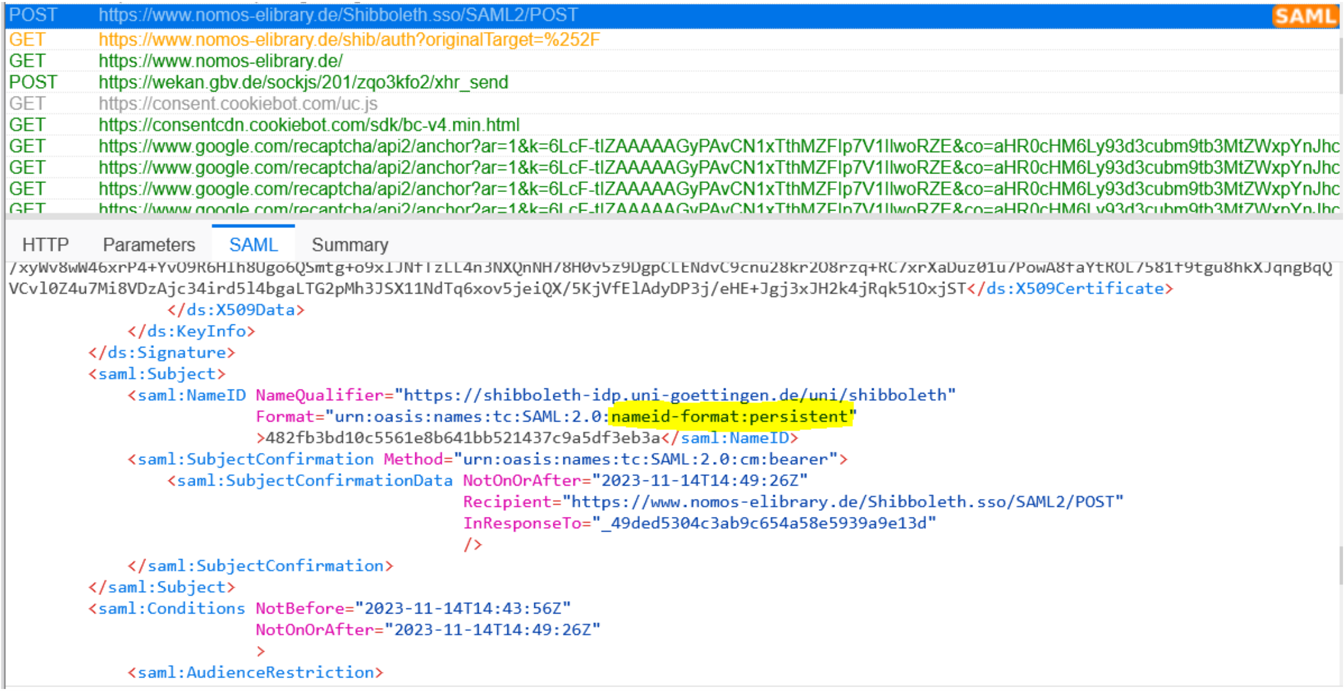
Pairwise persistent NameID (kein Attribut)

- Die SP einiger Anbieter fordern zusätzliche Informationen, die keine Attribute sind und daher den Nutzer*innen nicht in der Attributfreigabe der IdP anzeigt werden, von unserem IdP an.
- Die meisten Anbieter fordern eine transient NameID an, die pairwise ist und nicht persistent. Einige Anbieter fordern allerdings eine persistent NameID an.
- Genau wie „pairwise persistent“ Attribute kann auch die persistent NameID zur Wiedererkennung der Nutzer*innen verwendet werden.
- Das persistente Attribut Pairwise Id soll die persistent NameID ablösen (wie auch die eduPersonTargetedID). Als Attribut wird die Pairwise Id im Gegensatz zur NameID in der Attributfreigabe der Identity Provider aufgelistet.

Anbieter/Plattformen mit persistent NameId (Auswahl)

- Beck eLibrary (C.H. Beck eLibrary, Vahlen eLibrary), Plattform Nomos: https://www.beck-elibrary.de
- Classiques Garnier Numérique: https://classiques-garnier.com
- Duncker & Humblot: http://elibrary.duncker-humblot.de
- JSTOR: http://www.jstor.org
- Meiner eLibrary: https://meiner-elibrary.de
- Nomos eLibrary, Plattform Nomos: http://www.nomos-elibrary.de
- Thieme: z.B. https://eref.thieme.de, https://roempp.thieme.de (nicht thieme-connect.com)

SAML-tracer (Beispiel Nomos)



Zusammenfassung und Wunsch nach Austausch

- Bei der Anmeldung mit Shibboleth werden mitunter personenbezogene Daten vom IdP an den SP übermittelt.
  Manchmal sind diese für die Nutzer*innen sichtbar und manchmal nicht.
- Z.B. mit dem Firefox Add-on **SAML-tracer** lassen sich die übermittelten Informationen kontrollieren, auch wenn sie in der Attributfreigabe nicht angezeigt werden.
  (Bei verschlüsselter Übertragung ist dies allerdings nicht möglich.)
- Wir freuen uns über einen aktiven Austausch und sind unter der Mailadresse sso-team@sub.uni-goettingen.de zu erreichen.

SAML-tracer