

# Resource Summary Report

Generated by SPARC SAWG on Jul 29, 2026

## GROTOAP

RRID:SCR\_002631

Type: Tool

---

### Proper Citation

GROTOAP (RRID:SCR\_002631)

---

### Resource Information

**URL:** <http://ceon.pl/en/research/solutions/>

**Proper Citation:** GROTOAP (RRID:SCR\_002631)

**Description:** Test set for training and performance evaluation of page segmentation and zone classification algorithms. Contains input articles in a digital form and corresponding ground truth files. The test set is based on articles obtained from DOAJ database published under CC-BY license. The whole test set is available under the same license.

**Abbreviations:** GROTOAP

**Synonyms:** GROTOAP: GROund Truth for Open Access Publications, GROTOAP: GROTOAP: GROund Truth for Open Access Publications

**Resource Type:** data or information resource, data set, narrative resource, training material

**Keywords:** document analysis, page segmentation, zone classification, system evaluation, ground truth, text extraction, text mining, performance evaluation

**Funding:** National Centre for Research and Development Poland SP//I/1/ 77065/10

**Availability:** Free, Available for download, Freely available

**Resource Name:** GROTOAP

**Resource ID:** SCR\_002631

**Alternate IDs:** nlx\_156053

**Alternate URLs:** <http://www.force11.org/node/4758>

**License:** Creative Commons Attribution License

**Record Creation Time:** 20220129T080214+0000

**Record Last Update:** 20260729T044029+0000

---

## Ratings and Alerts

No rating or validation information has been found for GROTOAP.

No alerts have been found for GROTOAP.

---

## Data and Source Information

**Source:** [SciCrunch Registry](#)

---

## Usage and Citation Metrics

We have not found any literature mentions for this resource.