

keyword-extraction

March 9, 2024

```
[1]: import zipfile

with zipfile.ZipFile('SemEval2017.zip', 'r') as zip_ref:
    zip_ref.extractall('')

[3]: import spacy
from collections import Counter
import os

nlp = spacy.load("en_core_web_sm")

def extract_features(text):
    doc = nlp(text)
    features = {}

    for ent in doc.ents:
        features[ent.text] = {"entity_type": ent.label_}

    for token in doc:
        features.setdefault(token.text, {})[pos] = token.pos_

    return features

def extract_keywords(text, key_file):
    features = extract_features(text)

    potential_keywords = []
    if key_file:
        with open(key_file, 'r') as f:
            potential_keywords = [line.strip() for line in f.readlines()]

    candidate_keywords = []
    for key, value in features.items():
        if value.get("entity_type"):
            candidate_keywords.append(key)
        elif len(key.split()) > 1:
            if "NOUN" in value.get(pos, []) or "ADJ" in value.get(pos, []):
```

```

        candidate_keywords.append(key)

    if potential_keywords:
        word_counts = Counter(text.split())
        high_freq_keywords = [w for w in potential_keywords if word_counts[w] > 2]
        candidate_keywords.extend(high_freq_keywords)

    extracted_keywords = list(set(candidate_keywords))[:10]

    return extracted_keywords

def process_all_files(text_dir, key_dir):
    global extracted_keywords

    for filename in os.listdir(text_dir):
        text_file = os.path.join(text_dir, filename)
        key_file = os.path.join(key_dir, filename.replace(".txt", ".key"))

        if not os.path.exists(key_file):
            key_file = None

        extracted_keywords = extract_keywords(open(text_file).read(), key_file)
        print(f"File: {filename}")
        print(f"Extracted Keywords: {' '.join(extracted_keywords)}\n")

text_dir = "/content/SemEval2017/docsutf8"
key_dir = "/content/SemEval2017/keys"

process_all_files(text_dir, key_dir)

```

File: S0306437913000768.txt

Extracted Keywords:

File: S0305440314001927.txt

Extracted Keywords: Robson, two, Bates et al., SfM-MVS, James and Robson, 2012, Hess, 2a, 2010, Fig

File: S0021999112002847.txt

Extracted Keywords: DEA, two, 8, three, 9, Chebyshev, Fourier

File: S0167931712002936.txt

Extracted Keywords: 6, 60, 11, 13, DQD, gate, tens of pA., two QDs, 1c, Monte Carlo

File: S0039602899010493.txt

Extracted Keywords: He(1s, Hartree, AD, 4,6,23, DFT, 26, DOS, MIES, 5,6, 4-6

File: S0022311514007119.txt

Extracted Keywords: 1, 10, bcc Fe, 10,000, helium, 28, 29, First, 30a0×30a0×30a0, Fig

File: S0895611116300684.txt

Extracted Keywords: 2008, Oeltze, Wu, Wu et al., PSR, 2013, 2010, Preim, SA

File: S0370269304007701.txt

Extracted Keywords: 4,5, 1,2, 3, Quantum Field Theory, zero, half-integer

File: S2212667814001208.txt

Extracted Keywords: High Performance Computing, HPC, Cloud, Amazon, HPC Cloud, more than 20 percent

File: S0167931714004456.txt

Extracted Keywords: Elveflow (USA, SoftLithoBox, 29, Scientific Device Laboratory, 31, LOC, Microfluidic Innovations (USA, Canada, PDMS, 25

File: S0168583X14003929.txt

Extracted Keywords: 29.8, 1, 6, LMIS, Ga, Group III, 5, 4, Gallium, one

File: S0370269304009037.txt

Extracted Keywords: decay spectra, 3, QGP, two, 2, mT-)spectra, RHIC, PYTHIA

File: S0168365912006207.txt

Extracted Keywords: 92, 93, PAs

File: S0045782514001947.txt

Extracted Keywords: flow, HPD, 13, Edeling et al., 9, Highest Posterior-Density

File: S0022311515301069.txt

Extracted Keywords: 6, activation, 3, two, capsule, monitor, BODEX, one, fluence

File: S0301679X13003289.txt

Extracted Keywords: Transmission Electron Microscopy, two, TEM.Nanoparticle Tracking Analysis, 120, TEM, agglomerates, Nanoparticle Tracking Analysis, NTA

File: S2212667814001440.txt

Extracted Keywords: robot

File: S0370269302014880.txt

Extracted Keywords: CERN, 1, the previous 15 years, 3, Drell, QGP, 4-7, 2, 8, charmonia

File: S0010938X15301554.txt

Extracted Keywords: 1, Volta, One, 11, 11-22, corrosion, particles, second, AA2024, 23-25

File: S0377025714001931.txt

Extracted Keywords: four, 8, tack, Tack, 9

File: S0165212511000874.txt

Extracted Keywords: 1, 6, 3, 7,8, 2, 5, 4, ducts, first

File: S2212671612001163.txt

Extracted Keywords: Czech, Slovak

File: S0021999114002587.txt

Extracted Keywords: 1,2, complex micro geometries, 2, continuum fluid, CFD, Mach, Monte Carlo

File: S0167273812003025.txt

Extracted Keywords: RAMSI, Fe₃, XRD, second, EXAFS, La₄Ni₃-xFe_xO₁₀, La₄Ni₂FeO₁₀, 7.5, Ruddlesden

File: S0010938X15002085.txt

Extracted Keywords: 1, oxide, surface, 111 days, less than 1%, 4%, profile

File: S0749603615302184.txt

Extracted Keywords: DBR, 85%, GaN, first

File: S0098300413002720.txt

Extracted Keywords: one, Ripley, 2007, two, 2002, Multi-Layer Perceptron, Kotsiantis, 2009, 1996, Rojas

File: S0021999115003939.txt

Extracted Keywords: Darwish, 13, two, Section 3.2, node, Issa, TVD, Ubbink, 12

File: S0032386109007290.txt

Extracted Keywords: linear elastin-like AB di-block, VPGFG-(IPGVG)₄, 51, ELR, DSC, 52, 50, 49, TEM, first

File: S1877750313000240.txt

Extracted Keywords: 1, 21, LB, HemeLB, 22,23, Coveney

File: S221266781400080X.txt

Extracted Keywords: PAM, 60%, 55% and, 135.2mg, COD, FeSO₄, Al₂(SO₄)₃, FeSO₄

File: S0377025713001031.txt

Extracted Keywords: 1, 29, DoD, 30

File: S0079642515000705.txt

Extracted Keywords: film, Continuum, Ref, 39, Poxson, 228, 40, 42,266,267

File: S0003491613001516.txt

Extracted Keywords: Boltzmann, 1,2, nonabelian gauge theories, two, 3-8, recent years, four, 11-16, Complex Langevin, Langevin

File: S2212667812000524.txt

Extracted Keywords: Digital, library

File: S0009261412013838.txt

Extracted Keywords: 1, 6, 3, nanometre, graphene, fullerene, 4,5, AC-TEM

File: S030193221400144X.txt

Extracted Keywords: Alekseenko et al, film, LIF, 2014a, 2009, 2012, three

File: S0021999114008432.txt

Extracted Keywords: Wigner, VDS 0 V.
, Lres, 8, Section 4.3, 44, WTE, 47

File: S0370269304005829.txt

Extracted Keywords: 1, 3, 5×10^{-11} B., 10, about 170, 1015, MeV, 210, 3(b, 55

File: S0022311514008691.txt

Extracted Keywords: non-ODS, 14YWT, W, ODS, Y203, steels, 10,11, 12-15, between 9 and 14at.% Cr, 14at.%

File: S0375960113010839.txt

Extracted Keywords: 1, quantum, GT, 1(b, Fig

File: S0009261415000974.txt

Extracted Keywords: 1170K., 3, $V1-V2 \ln((\sigma_2^{-2})/(\sigma_1^{-1}))$, from Eq, second, seven

File: S0038092X15000559.txt

Extracted Keywords: 2008, Radiance, Jensen, 2001, Hachisuka et al, Hachisuka, photon, Monte Carlo

File: S1574119215000796.txt

Extracted Keywords: TDMA, MWSN, MAC

File: S0022311515300295.txt

Extracted Keywords: 1, around 300, a few hundred degrees, 2, 1000, Fukushima Daiichi

File: S1364815216302705.txt

Extracted Keywords: New York, 2014, Ferdaña, Gesch, DEM, 2009, 2012, SLAMM, Gilmer, Schmid

File: S0370269304006070.txt

Extracted Keywords: 23 45, lepton, 2, KamLAND, Third, Table 1, Ref, two, six, six-zero

File: S0370269304009335.txt

Extracted Keywords: eleven, $d > 4$, Minkowski, Hilbert, four

File: S2212667814000975.txt

Extracted Keywords: Jing-Jin-Ji, Yangtze River Delta

File: S0963869514000875.txt

Extracted Keywords: SH, 25, 18, 19, EMATs

File: S0098300413002951.txt

Extracted Keywords: Hitherto, three

File: S1359646214000165.txt

Extracted Keywords: PBE-GGA, Vanderbilt, 1, 2.75, 21, 22, DFT, 23, Brillouin, the Cambridge Serial Total Energy Package

File: S1010603013001809.txt

Extracted Keywords: Fe₃, RuO_x, 13, two, Cu⁺, CoO_x, IrO_x, co-catalysts, CoPi, one

File: S0021961413004321.txt

Extracted Keywords: 1-3, 5 to 300, 4-6, Smmechmix=XASmA+XBSmB, first

File: S0166361516300926.txt

Extracted Keywords: Enterprise Social Software, ERP, Chinese

File: S0022311515002391.txt

Extracted Keywords: powder, ODS, 50-80, DC, HIP, SPS, Joule, 17, mould, a few hours

File: S2212667813000774.txt

Extracted Keywords: EMO, Evolutionary Algorithms

File: S0167931713002487.txt

Extracted Keywords: wafers, Oxford Instruments OpAL, HP 4192A, HF CV, O₂, MBE, MOS, HfO₂, 250, Ge

File: S2352179114200056.txt

Extracted Keywords: 1, 6, 3, m², only half, 2, 4, 5

File: S0032386109007423.txt

Extracted Keywords: water, 2, IEC, three, PEMs, proton, Fig

File: S0378381215300297.txt

Extracted Keywords: 57, 56, EoS, two, kB. The Mie, decades, 55, 50-53, SAFT, LJ

File: S0009261408017028.txt

Extracted Keywords: two, 1560, 14, 8

File: S0370269304007798.txt

Extracted Keywords: quark, JPC, $q\bar{q}g$

File: S0378381215301291.txt

Extracted Keywords: 1, 1940, 3, 21-30, 2, the last ten years, 31, 36-41, first

File: S0022311515301653.txt

Extracted Keywords: US, one, Olander, UO2+x, + xCO
, fuel, Generation IV, 2, UCO, VHTR

File: S0168365913009036.txt

Extracted Keywords: LAP2, PDLAP1, incubated, PDLAPRG, 2, LCPRGPD, LA, peptide,
Method 1, liposome

File: S0370269304009086.txt

Extracted Keywords: collisions, K^* , Σ , N, $N^*(1440$

File: S2212671612000613.txt

Extracted Keywords: one, Canny, VOG, WIFI

File: S2212667814000756.txt

Extracted Keywords: 850, WGSR, three, hydrogen, 800

File: S0301010414003115.txt

Extracted Keywords: ANO, LR-CCSD, Abelian, 13, EOM, CCS, 6s5p4d3f1, B3LYP, CC3,
1s2s2p3s3p

File: S0038092X15003059.txt

Extracted Keywords: 2014b, 2, Yandt, two, 2.5, High Concentrator PV, 2012,
2014a, Fernández, first

File: S2212671612002351.txt

Extracted Keywords: prism, 7

File: S221450951530005X.txt

Extracted Keywords: 14cm, poly-blocks, approximately 1kg, the University of
Western Ontario, layer, 20×20×80cm³, four, Fig, 10min, LPF

File: S0375960112002885.txt

Extracted Keywords: cluster, several tens, First, 02

File: S0370269304008858.txt

Extracted Keywords: QCD, 10, 13, Ref, Schrödinger, BFKL, LLA, SYM, 17,
Heisenberg spin

File: S0021999112003579.txt

Extracted Keywords: LBB, 34, 10, two, Lagrange, Cxl, 9, Newton, Eq, first

File: S0370269303015478.txt

Extracted Keywords: 20, SM, 21

File: S0021999115003423.txt

Extracted Keywords: PDE, 2d, Section 2, Section 5, Section 4, Section 3

File: S002199911200068X.txt

Extracted Keywords: PDE, linear, Myocardial, thousands, 16-19, forty-million degrees, 13-15, DOF

File: S0370269304009530.txt

Extracted Keywords: lepton, particles, Monte Carlo, SUSY

File: S088523081530036X.txt

Extracted Keywords: .

File: S0010938X12001163.txt

Extracted Keywords: EDS, 20kV, 30kV., TEM, Focused Ion Beam, Pa., 1 m, FIB, 266, 2000FX

File: S0370157309002877.txt

Extracted Keywords: HRV, 3, 5, 2

File: S0021999116303291.txt

Extracted Keywords: PUT, NEMD, 9, 14, 10, 11, 13, DPD, 5, 8

File: S0888613X16301062.txt

Extracted Keywords: two, RO)BDDs, 9, PB, ProbLog, first

File: S030439911200040X.txt

Extracted Keywords: Canny, Thon, 23,24, 22, 2D, 1D, 23, 4,8,11, 12,14,18,20, PSD

File: S0377025714001682.txt

Extracted Keywords: 6, scattering, DLS, 240min, first, Fig

File: S0370269304009347.txt

Extracted Keywords: Wigner, 1f, Poisson, 10, 11, 1) $n = -\omega E n + 1 \sim (E) dE - n$

File: S0167931714003347.txt

Extracted Keywords: film, Brittle, TTC, 1c, 1%, 1b, fracture, strain, 1a, first

File: S0377221716300984.txt

Extracted Keywords: two, Permutation Flow-shop, three, Scheduling and Capacitated Vehicle Routing, framework

File: S0927025612000249.txt

Extracted Keywords: 1, two, late 1970s, 2, early 1990s, P91, 9-12%

File: S0021999115004301.txt

Extracted Keywords: Huynh, Galerkin, Lagrange, 23

File: S0021999113003422.txt

Extracted Keywords: IBM, 14, Peskin, Cartesian, oscillations, second, ISPM, Lagrangian, Maple

File: S0021999115008256.txt

Extracted Keywords: 2.3, 2.1, linear, Galerkin, 2.2, 9, Sec

File: S0165212511000862.txt

Extracted Keywords: 26, 24, Navier, 25

File: S0029549314002854.txt

Extracted Keywords: 33, 116, less than 10%, austenitic & ferritic stainless steel, 83, Fatigue, Three

File: S0254058415300766.txt

Extracted Keywords: ZnS, HMF, 1, un, 3, ZnTe, TiO₂, two, CdS, Half

File: S1361841516300822.txt

Extracted Keywords: 1, Table 1, Significance, Wilcoxon, two, between 3 and 7, F, 7 min, 7, T5

File: S0098300414002532.txt

Extracted Keywords: 463, 2, 2008, HRSC, Williams, about 91%, 2007, Tools for Graphics and Shapes, DEM, Steinwand

File: S0010938X12002508.txt

Extracted Keywords: Impressed Current Cathodic Protection, 1, 5-16years

File: S0010938X1500195X.txt

Extracted Keywords: -21S, 7,8, 1-3, oxidation, 1,4-6, alloys, 550

File: S1566253516300069.txt

Extracted Keywords: Konieczny, Pino-Perez, 67

File: S2212671612000121.txt

Extracted Keywords: quantum, three

File: S0370269302011838.txt

Extracted Keywords: RHIC, First, particle, SPS

File: S0038092X14000942.txt

Extracted Keywords: 273, 11, Table 3, 283, 5, 800W/m², 0.01kg, 400, 600, Three

File: S0168365913002848.txt

Extracted Keywords: 150 g, 10min, i.v., 4h, NB, Caliper Life Sciences,

Hopkinton, 5mg/kg, MA, 17

File: S074756321630348X.txt

Extracted Keywords: SNS, US\$17.4 billion, Australia, annually, Spiral, daily, November 2012, 2015, 16%, SNG

File: S0009261415008362.txt

Extracted Keywords: 20, 1, 21, 2L, 25, two, quantum, CCS, 22,23, 18

File: S0370269304007749.txt

Extracted Keywords: QCD, 9, 10, two, 1-3, quantum chromodynamics, NRQCD, 5, 4, 6,7

File: S0257897213004131.txt

Extracted Keywords: 7, the uncoated AZ31 magnesium, coating, Fig

File: S2212667814000550.txt

Extracted Keywords: Heilongjiang province

File: S107158191630074X.txt

Extracted Keywords: buffer, first

File: S0167273814004408.txt

Extracted Keywords: 180, YSZ, GDC, STF0

File: S004578251300176X.txt

Extracted Keywords: WC-Co, Co

File: S0378381215300674.txt

Extracted Keywords: CH₃,H₂O, 150, two, 129, CH₂,H₂O, 112, SAFT- SW, 82, H₂O

File: S0370269304009141.txt

Extracted Keywords: Collins, HERMES, CLAS, Q2, 5, DIS, 6-12, Refs

File: S0167931711005120.txt

Extracted Keywords: 3-6

File: S0031920113001222.txt

Extracted Keywords: Datong, 3, Liu, Changbaishan, Tengchong, Chinese, 2012b, 2009, Hainan, Hetland

File: S0032386110004039.txt

Extracted Keywords: 37, ROMP, 34, 39, ethyl dimethylamino, 35,36, 38, ROM

File: S0301932215002037.txt

Extracted Keywords: 92, 1997, first, waves, Sekoguchi, 2004, Sawai, 2.325

File: S0888613X16300767.txt

Extracted Keywords: MAP, CRG

File: S0266352X16301550.txt

Extracted Keywords: 15, 4,5, 6, two, 7,8-11, 2,3, 12-14

File: S0370269304009359.txt

Extracted Keywords: PP, Second, meson, SM, decays, B^-0 , K^-0 , $K0$, $Bd0(t) \rightarrow J/K^*$, 8

File: S000926141301539X.txt

Extracted Keywords: 37, first, TST, 38-40, second, OTS, Chesnavich, Letter, Hamiltonian, PES

File: S092702561300267X.txt

Extracted Keywords: Situ Adaptive Tabulation, 1, Thermo-Calc

File: S2212671612001497.txt

Extracted Keywords: the State Grid Corporation, PMS, galloping, 2009-2010, The State Grid Corporation

File: S0370269304009104.txt

Extracted Keywords: 15, 14, Infeld, cosmology, brane, Nordström AdS, Letter, Reissner

File: S0167931713002438.txt

Extracted Keywords: 15, two, 18, 17, electrons, al., 16, Bersuker

File: S0375960113004908.txt

Extracted Keywords: 20, Bi_2Te_3 , Faraday, ARPES, Hall, TIs, 1-3, TM, Bi_2Se_3 , 14-17

File: S0010938X15301268.txt

Extracted Keywords: two, hydrozincite, 9a, 9, NaCl, $Cu_{40}Zn$, 9b, Fig

File: S0021999115007238.txt

Extracted Keywords: 5 days, wave, Gaussian, 500km, 48,49,20, 4000km \times 4000km, 67km, 0.1Hz, the "Garden Sprinkler Effect, 150km

File: S2212667813001068.txt

Extracted Keywords: two, Timur-Coates, permeability, NMR, porosity

File: S0166218X1300348X.txt

Extracted Keywords: two, Max-linear, max-linear

File: S0301679X14003272.txt

Extracted Keywords: Fouvry et al., Hirsch, Neu, 18, 17, 2., two hundred, Fig

File: S0370269304007129.txt

Extracted Keywords: lepton, $\rightarrow e$, ETC, dipole

File: S003238610900086X.txt

Extracted Keywords: 1, 2, $q(=4 \sin$, 10, scattering, stem, Eq, 1A, 100, 2

File: S2212667814000732.txt

Extracted Keywords: recent six years, Caofeidian, China, Tangshan Bay, five

File: S221450951400014X.txt

Extracted Keywords: bitumen, greater than, 15%, 3, Marshall, particles, 2, 1%, glass, 20%

File: S221266781400121X.txt

Extracted Keywords: Polish, algorithm, English, Word Sense Disambiguation, Measure of Semantic Relatedness, Princeton WordNet

File: S0370269304008809.txt

Extracted Keywords: Weyl, Poissonian, two, Dyson, 8, one, 7

File: S0370269304009803.txt

Extracted Keywords: 6, 12) $+(p_1) - (p_2)$ 0, Q8, One, $\Gamma()=149.2\text{MeV}$, two, 5, 11) $ih_4 V(P P- PP,$, three

File: S0021999113002945.txt

Extracted Keywords: 1, $\|y\|_\omega$, 22, maximum-norm, 2, Inequality, 24

File: S2212667812000949.txt

Extracted Keywords:

File: S0301932214001499.txt

Extracted Keywords: Karabelas, Belt et al., 1991, 1997, Mathie, 2012, gas, Hewitt, 2010, Alekseenko et al.

File: S1877750313001269.txt

Extracted Keywords: HPC

File: S0370269304009220.txt

Extracted Keywords: One, 5, $U(1, 10, 11$, Lagrangian, Letter

File: S0010938X14000420.txt

Extracted Keywords: 15min, 6 m, 3 m, One, polished, surface, 99.5%, 5min, diamond, coupons, 3×5min

File: S0032386109010386.txt

Extracted Keywords: two

File: S0263822312001468.txt

Extracted Keywords: 1, 6, 80s and 90s of the last century, 5, 2, Hull

File: S0038092X14004824.txt
 Extracted Keywords: 1940, Belgium, Linke, WMO Task Group, DNI, Ångström and Rohde, WMO, Section 3.2, Ulmitz, 1959

File: S0045782513001448.txt
 Extracted Keywords: 1,2, CFL, 3, SPH, 4, the Courant-Friedrichs-Lewy, WCSPH

File: S0021999114008523.txt
 Extracted Keywords: 6, CA, Lockerby et al, 5, first

File: S221267161200217X.txt
 Extracted Keywords: Enumeration, first

File: S2212671612000741.txt
 Extracted Keywords: Kernel Extensional Matrix Matching, KEMM, Gale, 30%

File: S2212671612001709.txt
 Extracted Keywords: two, NC, CL

File: S0022311515002470.txt
 Extracted Keywords: 450, 4, C1, electrochemical, EDXD

File: S2212667812000792.txt
 Extracted Keywords: Bonita

File: S0301679X14000449.txt
 Extracted Keywords: 2×10^4 , 22, 11, 13, Vickers, RH, SiC, Fig

File: S2212667812000536.txt
 Extracted Keywords: three

File: S0370269304007257.txt
 Extracted Keywords: 23, H, two, 25, second, $e+e^-$, 2, 24

File: S0370269304009608.txt
 Extracted Keywords: 20, 21, NOMAD, , JETSET, half, 20%, 16

File: S0370269304012638.txt
 Extracted Keywords: 15, $SU(N)$, QED, Coulomb, F6, 13, two, D3, $SU(N-1) \times U(1)$, one

File: S0045782515002686.txt
 Extracted Keywords: IBM, 6, Peskin, 3, Newtonian, 5, 8, 4, Poiseuille, one

File: S0736585316300661.txt
 Extracted Keywords: Triplett, 0.3212, 2006, Supplementary Table S5, 0.3254, 2013, 0.3534 for HT

File: S0045782512002599.txt

Extracted Keywords: IETI, Poisson, linear, two, NURBS-, Non-Uniform Rational B-spline, Isogeometric Tearing and Interconnecting, IGA, FETI, NURBS

File: S2352179114200032.txt

Extracted Keywords: 1, 25, two, TCR, 02, gas, 12, first

File: S0379711213001653.txt

Extracted Keywords: 15, 14, 13, 18, FIRECAM, 19

File: S0038092X15001024.txt

Extracted Keywords: four additional morning, the entire year, 5h, the year, Markov, the summer, 5am, 5, 8, the four seasons

File: S0032386110001667.txt

Extracted Keywords: 67-69, 66, Schrock, 2

File: S0370269304006720.txt

Extracted Keywords: Forger and Gomes, 6, first, Grassmannians, 1-5, 13, two, Lüscher, Cartan, non-Abelian

File: S0167931712002699.txt

Extracted Keywords: ICP, 28, 25-27, silicon, first

File: S2212671612000431.txt

Extracted Keywords: Second, AdaBoost, symbian, recognition, LBP, First, ER

File: S0045782513000479.txt

Extracted Keywords: 20, 1, 21, two, Hamilton, 22,23, three, Jacobi, Osher, Sethian

File: S0925838814009669.txt

Extracted Keywords: Diffraction, 14, XRD, 25, 13, AA6061, AA5083, 500, SPS, 20h

File: S0370269304009074.txt

Extracted Keywords: 6, Friedman-Robertson-Walker, McVittie, Nordström, Einstein, Letter, Reissner, first, Schwarzschild

File: S0021999115008153.txt

Extracted Keywords: GPGPU

File: S0032386109005485.txt

Extracted Keywords: polymerization, CRP, 2, ATRP, 83, 90, 103, around 4, Fig

File: S0377221716303873.txt

Extracted Keywords: 3525, 8.75%, 0.53%, 25×25×12meter³, iron, 0.04%, 0.02%, first year, SiO₂, Dimitrakopoulos

File: S0021999113005846.txt

Extracted Keywords: Kelvin, f, forcing, damping, Chapters 9 and 10, 16

File: S0045782513000546.txt

Extracted Keywords: 14, 3,24, Galerkin, Lowrie, one, Euler

File: S0370269304009165.txt

Extracted Keywords: 20, Ref, 30, Letter, Hubble

File: S0370269303017222.txt

Extracted Keywords: 1, first, linear, Eq, $2\Delta M = -14C$, 29, one, Roper, q3

File: S1524070312000380.txt

Extracted Keywords: 3, two, Galerkin, CAD, Isogeometric, parameterized, NURBS

File: S2212667812000937.txt

Extracted Keywords: AI

File: S0167273814004548.txt

Extracted Keywords: Two, 11, seconds, 2a, 24, Fig

File: S0021999113004555.txt

Extracted Keywords: only one, HEVI, 1, Strang, 23, Ullrich, Rosenbrock, Newton, Jablonowski, Ascher

File: S0377025715000993.txt

Extracted Keywords: 1s-1, 1, about 4mPas, 10-4000s-1, 21, 15s-1, 4000s-1, ARES, 4, PEDOT

File: S0370157312000105.txt

Extracted Keywords: Hulse, today, Kip Thorne, Eddington, Kenneth Nordtvedt, PPN, Dicke, 2, the 1960s, the early 1970s

File: S2212671612001783.txt

Extracted Keywords: Ti-Al, intermetallic

File: S0379711215000223.txt

Extracted Keywords: 10-12, 13, 11, parameters, SCE, between 1% and 25%, e.g. Genetic Algorithm, Shuffled Complex Evolution, four, Fernandez-Pello

File: S037026930301801X.txt

Extracted Keywords: 14, two, 10-12,14Be, 12, first, Fig

File: S002199911300346X.txt

Extracted Keywords: 37, SGH, 10,24,25,3, Riemann, CCH, Lagrangian SGH, Wilkins, TOODY, Hallquist, many years

File: S221450951400031X.txt

Extracted Keywords: Electrotech, up to 5, Galvashield CC, Fosroc, ASR, Mott

MacDonald, 2005

File: S2212667814001166.txt

Extracted Keywords: 3d, max, LCD

File: S0377025715000051.txt

Extracted Keywords: 15, G , GP, FTMS, 9, G*(, Fourier Transform Mechanical Spectroscopy, 9,19,20, 15,16, Fourier

File: S1364815216303541.txt

Extracted Keywords: Utah, two, survey, three, Section 4.3, Utah Water Survey

File: S0045782512000266.txt

Extracted Keywords: 24-26, Firstly, second, 23

File: S0377025714002213.txt

Extracted Keywords: GP, SDG, second, 8, 9, third, first

File: S2214657115000052.txt

Extracted Keywords: 1, AC, 0.56%, 1-3, Finland, Only 4, 4, 0.7%, Fig

File: S2212667814000884.txt

Extracted Keywords: air

File: S0022311513011422.txt

Extracted Keywords: 6, Y2Ti2O7, Y2O3, ODS, 10, oxide, Ratti, Whittle, 9, one

File: S2212671612002181.txt

Extracted Keywords: 6, ELM

File: S0370269304009062.txt

Extracted Keywords: MBR, UHECR, two, 1019, IRB, Universe

File: S0370269304007567.txt

Extracted Keywords: 20, chamber, MAGBOLTZ, SciFi, 21, GARFIELD, 100 m

File: S0301010413002139.txt

Extracted Keywords: 15, 20, Hartree, 18, 10,16,17, 19

File: S0301932213000487.txt

Extracted Keywords: 15, 14, 17, wall, particle, one, Figs, Fig

File: S1631070514000954.txt

Extracted Keywords: 13, N=89, 90, half, 8, 1958, 103, 12, first, Ac

File: S2212667812000883.txt

Extracted Keywords: The 21st century, Chinese

File: S0021999113005603.txt
 Extracted Keywords: Mth, 6, velocity, 25, the core micro region, 0 yi hcs, 16, yi =0, 24, Nth

File: S2212671612001291.txt
 Extracted Keywords: 400.1, Al-Si, 2000

File: S0045782515001899.txt
 Extracted Keywords: DG, CFL, the Discontinuous Galerkin, 21-27, recent years

File: S0010938X1530161X.txt
 Extracted Keywords: only one, sample

File: S0032386108009397.txt
 Extracted Keywords: 66-68, NL, PDMS, NS, 69, two, 9a, Mg, elastomers, 70

File: S0009261414000372.txt
 Extracted Keywords: 1, QED, 850, two, 2-4, 8-10, 800

File: S0254058414000662.txt
 Extracted Keywords: Second, ylideC2/C6, 33, Fukui, 3, ylideC7, HOMOylide, second, ylidepyrene, ylide

File: S221266781400149X.txt
 Extracted Keywords: four, Recursive Least Squares, RLS, Sample Matrix Inversion, CGM, beamforming, LMS, Least Mean Squares, SMI

File: S0167931712003905.txt
 Extracted Keywords: 20, 4.26, quantum, 50, 1.4×10^5 A, 1.7×10^5 , 80, FIB

File: S0009261412006513.txt
 Extracted Keywords: XAS, XES, One, two, water, recent years, 1-9

File: S0370269302012492.txt
 Extracted Keywords: Coulomb, 23, 13, 11, $(r, \text{six}, 21-23, t) = (r)_{ei} t$, four, 24

File: S2212667814000124.txt
 Extracted Keywords: EM, PX

File: S1364815216303061.txt
 Extracted Keywords: 1998, Hammer, Powell, Kadlec, Zhang, WATFLOOD, 2008, Wallace, two, 2012

File: S2212671612000704.txt
 Extracted Keywords:

File: S2212667814000987.txt
 Extracted Keywords: DC, AC, MSTU

File: S0927025614003322.txt

Extracted Keywords: API, 14, AFLOWLIB, 19-31

File: S0045782515003680.txt

Extracted Keywords: today, 1, 11, 2-8, CAD, 12-15

File: S0021999115008372.txt

Extracted Keywords: more than twenty years ago, Dietachmayer, 14, 26, many decades, meshes, the Monge-Ampère, Monge-Ampère, mesh, first

File: S2212667814000951.txt

Extracted Keywords: Expression, Force and Qi, China, Nature and Man in One, Chinese

File: S0032386109003991.txt

Extracted Keywords: polymerization, first, PS

File: S0022311514006722.txt

Extracted Keywords: oxidation, 1-5

File: S0009261415001517.txt

Extracted Keywords: one, two, 7,8

File: S2214509515300103.txt

Extracted Keywords: concrete

File: S0257897213003563.txt

Extracted Keywords: 15, 159, A0, Fabry, TiO₂, Adachi, 1.5eV to 3.5eV, SEM, 4, 70

File: S0010482516301810.txt

Extracted Keywords: at least one day, Amira, Stratasys Ltd., 3D-DSA, Eden Prairie, OBJET30 Pro, FEI, STL, 0.028mm, Burlington

File: S0022311514000919.txt

Extracted Keywords: second, Section 4.3, solidification, liquid, first, Fig

File: S0010938X15002954.txt

Extracted Keywords: 11,12,37, Tang, 34, 33, 2D, Newman, 1D, Ryan et al, 35, current

File: S0370269304009013.txt

Extracted Keywords: 14, 13, two, Ref, 17, Pauli, reaction, one, 15,16, 12

File: S2212667814001348.txt

Extracted Keywords: Matlab, the ETD Runge-Kutta 4, The Exponential Time Differencing (ETD, Kawahara

File: S0167273813005298.txt

Extracted Keywords: 20, 23, Ta, 200, 300, 32f, 500, 8c, 400, 100

File: S221267161200220X.txt

Extracted Keywords: Foreign Direct Investment, China, FDI, the next several years, GARCH-M

File: S2212671612002338.txt

Extracted Keywords: only one, Otsu, thresholding

File: S2212667814001361.txt

Extracted Keywords: three, un, quality

File: S0032386109005357.txt

Extracted Keywords: ROP, 29, 1,2,4, 2(a, 1,4, PLLA, poly(butylene, 2(d, hydroxyl, 27

File: S037026930400721X.txt

Extracted Keywords: Yukawa, MSSM, 6, 23

File: S0370269304007208.txt

Extracted Keywords: Coulomb, 1, 3) $w--F aFa =0$, Abelian, 6) $Leff=-14F aFa +M2-F aFa$, two, $F a= A a- A a+gfab cA bA c$, 2, 5, 4

File: S0968432814000250.txt

Extracted Keywords: 2008, De Winter et, 3, 2007, Knott, Hennig, Vihinen et al., Bock, 2012, 2009

File: S0039602899010869.txt

Extracted Keywords: 32, Hamaker, 8, the van der Waals, three, Si, 27, first

File: S037026930400680X.txt

Extracted Keywords: Yukawa, rhd neutrino masses, rhd neutrino masses
MR 1010-1015

File: S0022311514006849.txt

Extracted Keywords: 13, Willis, fourth, Geng et al, two, six, 8, Govers, 12, 6

File: S0370269304008421.txt

Extracted Keywords: Hamiltonians, 21, 12) $Pc \theta(A-\Delta)e-(\cos^2 rel+1)/2$, 5,22, neutrino, Sun, three, $x>0$, $A \Delta$, NSI

File: S1875952116300209.txt

Extracted Keywords: 20, EMDB, French, English, Film Stim, the Film Stim, the Emotional Movie Database, 19

File: S025405841530136X.txt

Extracted Keywords: Al, inoculants, grain, B. Moreover, 4, Nb+B, Nb, Al-Si,

first

File: S0022311513001165.txt

Extracted Keywords: first, ionization, W, -7.86eV, Larmor, two, FWupstrem, W.
With, second, divertor

File: S0370269301015222.txt

Extracted Keywords: 6, Λ , two, FRW, 1-3, $\Omega\Lambda$, 5, 4, about one third, ΩM

File: S0165212515000931.txt

Extracted Keywords: several two, 3, two, Millar, 2, 4, three, the 1960s,
Shenderov, 4-6

File: S1877750311000676.txt

Extracted Keywords: One, 2D, 7, 2, VISSIM, simulation, Fig

File: S2212667812000664.txt

Extracted Keywords: recognition, PCA, PRS, analysis(PCA

File: S0021999115007895.txt

Extracted Keywords: one, Hamilton, 1,19, Lagrange, Noether, Lagrangian,
Hamiltonian, Euler

File: S0370269304007816.txt

Extracted Keywords: two, three, Bargmann, Calogero

File: S0022311514006941.txt

Extracted Keywords: the Australian Nuclear Science and Technology Organisation,
Table 1, Sydney, x=1, B203, 4, glass, 6, Carter, 5

File: S221267161400105X.txt

Extracted Keywords: Microsoft Kinect, between two, 1, Ten, two, second, SVM,
first, SURF

File: S0997754612001318.txt

Extracted Keywords: Stuart, Lingwood, 2, Bödewadt, von Kármán, 7,10,11, Dhanak,
BEK, 8, 4

File: S2212667812000780.txt

Extracted Keywords: learning, teaching, china

File: S0254058415304235.txt

Extracted Keywords: Mossbauer, 25, 23, Zener, 24, iron

File: S0301010413004096.txt

Extracted Keywords: INS, NPD, 5K, MOF, H2, D2, first

File: S1386505616301769.txt

Extracted Keywords: two, Kibera, EMR

File: S2212671612001692.txt

Extracted Keywords: recent years, three

File: S0022311515303901.txt

Extracted Keywords: 15, 14, 12,13, CH4, graphite, 16, CO, Magnox, 12, first

File: S0032386114008428.txt

Extracted Keywords: HEMA, 52, 2, BA, MMA, 53,54, poly(butyl

File: S0370269304006768.txt

Extracted Keywords: 8, 2

File: S0370269304009025.txt

Extracted Keywords: MIT, Meson Coupling, Vvq, 8, QMC, 10,11, 9, 12,
30)i · -m-Vqs- 0Vqv QMq(x)=0

File: S092583881302834X.txt

Extracted Keywords: Fujita et al., first, 30%, 16%, 18, just 4%, up to 24%, 17,
MGs, MG

File: S2212667813000762.txt

Extracted Keywords: 20, Yinchuan Plain, about 6 acres

File: S136481521630305X.txt

Extracted Keywords: two, daily, N20, second, SB, three, MSE, N, first, UK

File: S1877750315000460.txt

Extracted Keywords: HemeLB, FabHemeLB, FabSim, 27, one

File: S1877750315000575.txt

Extracted Keywords: One

File: S0022311511010014.txt

Extracted Keywords: Pu, PuGa, Ga

File: S0045782513001473.txt

Extracted Keywords: 3, 26,14, 25, Oij, Schweitz, 16, Eq, 32

File: S0168874X1630049X.txt

Extracted Keywords: 14, Grassl, 2(b, 5, CDPM2, 16, Fig

File: S0370269304009657.txt

Extracted Keywords: 0-, Salpeter, 18, 17, Letter, 16

File: S0098300413002124.txt

Extracted Keywords: Digital Elevation Model, 3DWF, Earth, World Geodetic System,

Fukushima, WGS84, 2008, 56S, ENU, 60N

File: S0045782515001231.txt

Extracted Keywords: 1, 1,2, 3, 11, 5, BEM, CAD, 4, Isogeometric, 6,7

File: S0963869515000572.txt

Extracted Keywords: 1, propagation, wave, 13, SH, Fig

File: S0370269304009116.txt

Extracted Keywords: 3, two, Salam-Sezgin (Minkowski) $4 \times S^2$ ground state, 2, Poincaré, four, one

File: S0370269304009049.txt

Extracted Keywords: 1, QCD, 6, 3, One, Ref, two, quantum chromodynamics, 2, Hamiltonian

File: S2212667814001397.txt

Extracted Keywords: Journal Influence Score, This Influence Score, the Journal Influence Score, the SCImago Journal Score

File: S0375960115005630.txt

Extracted Keywords: plasma, electrons, Section 2, Stern

File: S0022311514001640.txt

Extracted Keywords: vapour, 1, 2100, 25, two, carbon, SiC, TRISO, 24, Fig

File: S0011227515001216.txt

Extracted Keywords: 20,000, 0.000968, 2, 3.57×10^{-6} , Eq, Hills et al, 1.7, RRR, 0.0328, 8

File: S2212667814001294.txt

Extracted Keywords: Role-Based Access Control, KM, RBAC, Knowledge Management System, Knowledge Management, KMS, the past decade

File: S2212667814001464.txt

Extracted Keywords: one, first

File: S0079642514000887.txt

Extracted Keywords: Goteborg, AFM, two, STM, 58-61, 83, 72, 27,76,77, Nanofactory Instruments AB, TEM

File: S0167844214000652.txt

Extracted Keywords: two, 3, micro-cracks, Fig

File: S0301932213001985.txt

Extracted Keywords: CLSVOF, 5th, WENO, 2nd, Xiao, three, 2012, VOF, one, LS

File: S0021999115008207.txt

Extracted Keywords: 21, 34, two, 50, 40, 18, 36

File: S2212667814000045.txt

Extracted Keywords: 8-12%

File: S0009261413006738.txt

Extracted Keywords: 14, 22,23, TSH, 24, PES

File: S0010938X14002157.txt

Extracted Keywords: 20, 15, Section 5.6, surfactants, 14,15,30,106,107, CI, linear, NPE, 106, polyoxyethylene sorbitan

File: S0370269304008974.txt

Extracted Keywords: OPE, JW, QCD, two, SDO, , 28, Secondly, 17, First

File: S0022311514009271.txt

Extracted Keywords: 1, two, Arima, Morelon, cation, Grimes, Catlow, Morl

File: S037026930400930X.txt

Extracted Keywords: Conrad, 34, Gaussian, 35, Highland, Hill, Fsens, Feldman, 27, Table 2

File: S0301010415300355.txt

Extracted Keywords: 1, AOH, about 2.7eV, recent years, 30-32, AcH, BAH, Fig

File: S0029549314002970.txt

Extracted Keywords: six, 1st, 0.1MPa, the 1st, Na

File: S0098300414000259.txt

Extracted Keywords: Apache Pig, Hadoop, Java, DAG, Mayer, 2013, Java MapReduce, Latin, Pig, MapReduce

File: S0301932214001931.txt

Extracted Keywords: Sommerfeld, particles, 2013, particle, Doppler, 2012, Mallouppas, LES, van Wachem, 2002

File: S2212667814001488.txt

Extracted Keywords: Java, Scripthon

File: S0957417416302561.txt

Extracted Keywords: Vujaklija, DFT, Park & Lee, two, Kuiken et al., 2014)11<https://www.coaptengineering.com/>., 2015, 2016, EMG, COAPT

File: S2212667812000895.txt

Extracted Keywords:

File: S0963869514000954.txt

Extracted Keywords: maxima, GA, 19

File: S0957417416303773.txt

Extracted Keywords: the Principal Component Analysis, the recent years, Jolliffe, 2002, PCA, first

File: S0927025614007137.txt

Extracted Keywords: 1,2, 14, 3, silica, 13, two, second, 5, 15-17, 4

File: S0370269304009189.txt

Extracted Keywords: 15, q4, 14, quark, 13, pentaquark, 4]0, 4, 16, 1/2

File: S0927025615006357.txt

Extracted Keywords: UMAT, LSHR, second, XFEM, RVE, third, first

File: S0010938X13005945.txt

Extracted Keywords: 200 m, 1-4N, Ra, eight, roughness, 6mm, linear, coating, Revetest, H-270

File: S002231151500032X.txt

Extracted Keywords: Polatidis et al, Petigny et al., 1, One, Synchrotron X-Ray Diffraction, 1-7

File: S0045782511003823.txt

Extracted Keywords: NURBS, 2002, Laursen, 25

File: S0032386109004996.txt

Extracted Keywords: SPM, 10, AFM, 15-18, polymer, one, crystal, 19

File: S0167931712003012.txt

Extracted Keywords: between 800 and 2000, between 30 and 325nm, SoC, two, up to 330, 12, IM, Irresistible Materials, IM-HM11-02, IM-HM11-01

File: S0370269304008998.txt

Extracted Keywords: LCDA, QCD, 14, 21, 10, 07, two, 11, NNLO, TiI

File: S0168365913004975.txt

Extracted Keywords: 25%, 24h, 45%, 1,6-hexanediamine, 1.0, 30%, 30ml, 50ml, 314mg, 18

File: S0254058415001212.txt

Extracted Keywords: 1, 5, TiAl, 2, titanium

File: S0032386110001254.txt

Extracted Keywords: 15, 1, 1-Alkyl-3-methylimidazolium, 21, One, 25, 26, 3-(2, 22,23, ATRP

File: S0022311515300830.txt

Extracted Keywords: 20, 23-114, 99.95, 1088.8, Bi2UO6, XRD, 7 ppm, two, Zn, 50

File: S2212667814000069.txt

Extracted Keywords: first

File: S000926141500651X.txt

Extracted Keywords: AA, 1, 3, two, second, DBS, 2, DSBs, 18, 3b

File: S0377221716304258.txt

Extracted Keywords: Bodrog, 1998, Meini, Breuer, Kronecker, 2008, & Lohmann, Bini, Casale, Zhang, & Smirni, Horváth & Okamura

File: S003238610801080X.txt

Extracted Keywords: microscopy, two, microscopes, one, first

File: S0022311515300477.txt

Extracted Keywords: TEM, nanometre

File: S0142061516308079.txt

Extracted Keywords: RT, 37, 34, 35, 36

File: S2212667814000331.txt

Extracted Keywords: Lastly, PT, Firstly, Secondly

File: S002199911500025X.txt

Extracted Keywords: two

File: S1746809416300933.txt

Extracted Keywords: ObjectiveElectrically, 900pps, linear, between 0.7 and 1.7ms, 500pps, 2ms, EEG, CI, 0.7

File: S2212667812000822.txt

Extracted Keywords: Ka la OK, WTO, China, Chinese, more than a century, Ecolinguistics

File: S0377025714000317.txt

Extracted Keywords: Denier, 9c, Robin, solutions, 10a, 13, 9a, Newtonian, Hewitt, ω 1

File: S0370269302014892.txt

Extracted Keywords: 25, Lepage, 13, two, 8, formula, 9, 24, Brodsky

File: S187775031300077X.txt

Extracted Keywords: 15, PDE, fourteen, Partial Differential Equations, about 2500cm², Liley, 18, Ordinary Differential Equations, 17, 16

File: S2212671612002120.txt

Extracted Keywords: Firstly, second

File: S0165168416300603.txt
 Extracted Keywords: 41, first, 10, TDOA, 39, SRC, two, $z_a = \{m(t, \text{GCC-PHAT}, 5$

File: S2212671612000637.txt
 Extracted Keywords: Hamiltonian

File: S0370269304009268.txt
 Extracted Keywords: Higgs, decay $H \rightarrow$, 1, SM, up to, QCD, One, the Standard Model (SM, 200, LEP

File: S0305054816300867.txt
 Extracted Keywords: 1, 3, linear, two, Louveaux, 26, Bender, Jensen, second, Batun

File: S0166218X14003011.txt
 Extracted Keywords: Gregory, Wielandt, Dulmage-Mendelsohn, max-algebraic, one, Kim, Schwarz

File: S0045782514004812.txt
 Extracted Keywords: 0, Hamiltonian, first, Δt_n

File: S2212671612001618.txt
 Extracted Keywords: LMI

File: S096386951400070X.txt
 Extracted Keywords: up to, SNR, 3.7, ECT, NERSE, three

File: S0022311513011951.txt
 Extracted Keywords: 17, Frenkel, ODS

File: S0370269304008305.txt
 Extracted Keywords: CCFM, Kwieciński, HERA

File: S0022311513010313.txt
 Extracted Keywords: Avus Metals & Plastics Ltd., 18/8, 15h, Acros Organics, 3, Calumite, Fenton Packaging Ltd., PCM, PVC, Romar Workwear Ltd.

File: S2214657115000155.txt
 Extracted Keywords: 6, 11, 13, microCT, Hot Isostatic Pressing (HIP, 8, 12, 0.01%, 0.2%, MicroCT

File: S0009261415002730.txt
 Extracted Keywords: Two, Second, 33, 5A, 5, 5C, three, MD, one, First

File: S0377025714000135.txt
 Extracted Keywords: Weissenberg, VE, 37, 3,6, DoD, 4, 38, Clasen, jetting

File: S0021999115001412.txt

Extracted Keywords: 37,33,32,41, non-Newtonian, 46,45,32

File: S167420011300196X.txt

Extracted Keywords: 1, S2, 3, only S4, 2006, Henry, S3, 2, Eg, Hanna & Nozik

File: S0370269304009979.txt

Extracted Keywords: 6, 1, six, 3/2, Sundrum, 2, 8, Standard Model, 1/2, graviton

File: S1364815216303243.txt

Extracted Keywords: Tsai, 2D, 2004, Hénnonin, LISFLOOD-FP, two, Bates et al.,
MIKE FLOOD, Hunter et al., 2012

File: S0021999113005652.txt

Extracted Keywords: Six degrees, two, water, second, Schmidt, 1 to 8, three,
DNS, gas, Reynolds

File: S0010938X13002187.txt

Extracted Keywords: one, 1, two, TGA, 100h, 650, 600, 1000h

File: S0168365913009589.txt

Extracted Keywords: US, 15, 14, acoustic, 16

File: S037596011300741X.txt

Extracted Keywords: 6, 269, some twenty years ago, Holland s, Miller, 3,
Hamilton, Schrödinger, 2, de Gosson

File: S0370269302013412.txt

Extracted Keywords: Einstein's, the Hoyle-Narlikar, $z>1.7$.We, $z>1.7$

File: S0168365912007560.txt

Extracted Keywords: pHrodo, S4G, LN, antigen, pHrodo Red, MD, CD11b, 24h, Ova,
FACS

File: S2212667813000610.txt

Extracted Keywords: Turkish, management

File: S0375960113006725.txt

Extracted Keywords: 6×10^{12} , 1, 2.5×10^{-23})-(2.5×10^{-21} , OH 18, Eq, 10 km, 21 cm, a
few tenths, 109, about 1014-1015

File: S2212671612000686.txt

Extracted Keywords: 2008, 2011, the first three quarters, Chinese, the fourth
quarter

File: S0045782515002418.txt

Extracted Keywords: Huynh, nodal, 2, 1D

File: S0021999113002362.txt

Extracted Keywords: Coulomb, 10, 11, Pauli, the Fermi-Dirac, Monte Carlo

File: S0003491615001955.txt

Extracted Keywords: 14-18, MIT, 13, Casimir, nucleon, Maxwell, one, first, 19

File: S0370269304009177.txt

Extracted Keywords: Higgs, two, solution, SUSY, S1, A5

File: S1359028614000989.txt

Extracted Keywords:

File: S0045782512003234.txt

Extracted Keywords: parameters.1, Section 3, Section 6, Section 5, Section 2, Hellmich, Section 4, Section 8), 30, Pivonka et al

File: S0032386109001463.txt

Extracted Keywords: Gemma, 1, Monte Carlo Simulation, two, only three, three, ABC, Dotera, 30, 31

File: S0370269304009232.txt

Extracted Keywords: quantum, singularity, Big Rip, Barrow

File: S0021999114007396.txt

Extracted Keywords: two, Galerkin

File: S004578251400334X.txt

Extracted Keywords: 17, meshes, 18, Reynolds

File: S0167273813006735.txt

Extracted Keywords: one, two, LSC, YSZ, IEDP, La0.6Sr0.4CoO3-, films, oxygen, yttria stabilized zirconia, EIS

File: S0263822312000657.txt

Extracted Keywords: 1, Suresh, Mortensen, 2, Functionally Graded Materials, Japan, one, the 1980s

File: S221267161200176X.txt

Extracted Keywords: HDMR, SVR

File: S0021999115003459.txt

Extracted Keywords: Bernoulli, 12,13, linear, BEM, First

File: S0021961415003821.txt

Extracted Keywords: 122, , six, 15-6, alkanes, 125, 118, 132, three, n-C15H32

File: S0021999115000546.txt

Extracted Keywords: 6, Allen-Cahn, Carn, One, 10, two, monotonically, 8, three, 9

File: S0393044012000198.txt
Extracted Keywords: Newman, 22, Minkowski, zero, 17, 19-21, Maxwell, RemarkThe

File: S0022311515002354.txt
Extracted Keywords: 5,41,42, 13, DHC, Hydrides, hydride

File: S0377221716302259.txt
Extracted Keywords: 2011, 1, Chang et al., two, 2, 2012, Variable Returns to Scale, Mahlberg & Sahoo

File: S2212667814000264.txt
Extracted Keywords: two, three, vAcademia

File: S221266781300018X.txt
Extracted Keywords: Amodel, The Linear Temporal First-Order Logic

File: S0167931713005042.txt
Extracted Keywords: 1×1cm², EBL, 1, 225, Advanced Diamond Technologies Ltd., RIE, SC-1, 2 m, HSQ, 17

File: S0963869514001078.txt
Extracted Keywords: 15, Lorentz, 18, 770

File: S2212671612000698.txt
Extracted Keywords: Obstacle, classification

File: S0045782512002678.txt
Extracted Keywords: BDDC, linear, 13, 12,13,16,29, three, 29, Schwarz, 12, 16

File: S2212667812000032.txt
Extracted Keywords: AOP, the Bank Account System, three, First, first

File: S0167931714000203.txt
Extracted Keywords:

File: S0167931713004991.txt
Extracted Keywords: PB, between 10 and 50, 19, QSF6, 20min, O₂, three, between 0 and 20sccm, four, Table 1

File: S1361841516300342.txt
Extracted Keywords: 1998, Toledo, Rivest-Henault, 1997, Curwen et al., 1994, Jomier et al., 2012, 2009, Condensation

File: S2212671612002375.txt
Extracted Keywords: Jena, Power Grid

File: S0045782514001492.txt

Extracted Keywords: 43, 45, Betsch, Chawla, Lagrange, 46, 44, Hesch, Laursen, 47

File: S0010938X13003818.txt

Extracted Keywords: 6009, $1 \pm 0.05s$, 50Hz, between 0.5 and 1023Hz, Rm, 1000, 1023Hz, 9, 0.5Hz, NI

File: S037877531001949X.txt

Extracted Keywords: Teflon, 34,35,36, coating, 35, 37,38,39, The Substrate Induced Coagulation, Li, SIC, 36

File: S016793171300244X.txt

Extracted Keywords: 6, SiO₂, two, 1-3, 5, 4, three, Si, one

File: S0032386109006612.txt

Extracted Keywords: 36,37, SWNTs, 3, One, 1,26, HiPco, 28-30, 27, CVD, 31

File: S0011227515000648.txt

Extracted Keywords: LNT, 1/#2, approximately half, 88, 110, 5/#6

File: S0168365913008766.txt

Extracted Keywords: IFN, TLR, 27, 30, CpG, one

File: S0032386108010392.txt

Extracted Keywords: 20, 21, Poisson, Tabor, 9

File: S2212671612002302.txt

Extracted Keywords:

File: S1010603009002676.txt

Extracted Keywords: 25, 26, SPA, TPA, TPF, one-half, RhB, 24

File: S0888327016302333.txt

Extracted Keywords: 10-12, linear, 8, 9, just two or three, Volterra

File: S0022311514005480.txt

Extracted Keywords: Parise et al., 1, 3, 34, 0.1GPa, second, 2, three, fracture, 31

File: S0963869514001066.txt

Extracted Keywords: the Total Focusing Method, TFM, Linear, FMC, xT, Hilbert, zero, by:(10)ITFM(x, 3a, Fig

File: S0022311515301963.txt

Extracted Keywords: 13, 4,5,31, 3.9, 2.4, 1.7, Schottky, 7.04-12.92, Morelon, Jackson, 12

File: S1566253516300252.txt

Extracted Keywords: 42, Euclidean

File: S0377221716301904.txt

Extracted Keywords: one, two, first

File: S0370269304009293.txt

Extracted Keywords: OPE, QCD, 3, EUS(r, $\Delta V(r)=V_0(r)-V_S(r, r\rightarrow)$, Wilson, $r \wedge QCD-1$

File: S092702561300760X.txt

Extracted Keywords: particle, Reynolds, MD

File: S0098300412001793.txt

Extracted Keywords: MATLAB, Brandelik, 2004, CALCMIN, AMPH, 1992, 1994, 2009, MINERAL, Sturm

File: S0370269304008780.txt

Extracted Keywords: 15, NJL, CFL, 2SC, Ref, RKH, Taylor, Nature, 16

File: S0045782512003428.txt

Extracted Keywords: CAG, NS-FEM, 1, 2,28, second, 8, SFEM, Lagrange, 16, W2

File: S0021961414003255.txt

Extracted Keywords: one, three, XRD, Raman

File: S0029549313003439.txt

Extracted Keywords: Oak Ridge National Laboratories, OECD-NEA, two, TSUNAMI, D4, three, the Organization for Economic Cooperation and Development - Nuclear Energy Agency, Uncertainty Analysis in Modelling, reactor

File: S0003491615000433.txt

Extracted Keywords: 21,22, quantum, 4 decades, 13,14,10, 4,23, C60, 2, 15,16

File: S0301010414003516.txt

Extracted Keywords: 5cm \times 5cm, Minispectros, 1mm, Aerre Machines, Kurt Lesker, Dupont-Teijin, DNTT, TPGDA, substrates, 25m

File: S0021999113006955.txt

Extracted Keywords: P, solution

File: S0167931713006904.txt

Extracted Keywords: TSV, Tafel, ECD, LSM

File: S0370269304006161.txt

Extracted Keywords: Raghavan, 13, e-(e+, 40K*., 40Ar, Gamow-Teller, Bhattacharya, MeV, Bueno, MEMA

File: S0010938X15003261.txt

Extracted Keywords: Kelvin, SKPFM, Volta, as high as 10's, metal, 14,15,19, Ecorr, 10,15,18,20-22, 18, EWF

File: S0167273815004130.txt
Extracted Keywords: 74, XRD, 850, JEOL, 1300, 5%, TGA, 1h, four, STOE Win XPOW

File: S0045782514000607.txt
Extracted Keywords: PL, convergence, 1, 2

File: S0021999114007876.txt
Extracted Keywords: RTE, DOT, FEM, Monte Carlo

File: S0032386109001712.txt
Extracted Keywords: 53, polymer, Mayer, crystallization, PVA, 51,52, Muller

File: S0370157314001318.txt
Extracted Keywords: 12, 10, 11

File: S1687850714000405.txt
Extracted Keywords: Xylanases, Kantelinen, Beck & Scoot, Ho, 1985, 1992, 1994, Jurasek, Buchert, & Puls, Walsh, Power, & Headon

File: S0168365913001521.txt
Extracted Keywords: 5h, TSI, Inc., silica precursor, MO, Sigma-Aldrich Corp., APTES, 80, MN, TEOS, 18

File: S0167273811005091.txt
Extracted Keywords: 12, FIB

File: S0370269304008494.txt
Extracted Keywords: 2, QCD, One, the meson, about 30%, 5,6, about one

File: S0011227514002136.txt
Extracted Keywords: 17, pill, temperature, Fig

File: S0010938X15301189.txt
Extracted Keywords: 74%, 1-5, 8%, 16%, Pressurised Water Reactors, 600

File: S1574119211001544.txt
Extracted Keywords: Firstly, two, nodes, 3.3, Secondly, node misbehaviours, Section

File: S0038092X11004129.txt
Extracted Keywords: 7%, 15%, approximately 8%, Table 2, 50, the 15%

File: S2352179115300041.txt
Extracted Keywords: 300-400, 10, KSTAR, 1-3, 350, 200-275, ITER, carbon, TCV

File: S0304399111001811.txt
Extracted Keywords: Schrödinger, zero, Bessel, vortices, plane

File: S0029549314001551.txt

Extracted Keywords: neutron, 30cm, core, sodium, 800pcm, 60cm, 2013, plenum, Fig, al.

File: S0375960115004120.txt

Extracted Keywords: vacuum, Poissonian, quantum field theory, Casimir, Lamb, particle, , zero

File: S0022311515002664.txt

Extracted Keywords: 1, 3, oxide, 2, , 23-26, 29, ± 30 , Table 1

File: S2212667812000810.txt

Extracted Keywords:

File: S0370269304007439.txt

Extracted Keywords: QCD, Gribov-Regge, 23, NEXUS, two, pomeron, 22,23, one, 3.97, Letter

File: S0079642514000784.txt

Extracted Keywords: N₂, quadrupole, CH₄, dipole moment, Li et al, CO₂, 91, SO₂, 90, H₂O. Thus

File: S0038092X15001681.txt

Extracted Keywords: 10, system, two, 5(b, 873, 864, 50kWp, 18 modules, 40kWp, one

File: S0031920113000708.txt

Extracted Keywords: Jerk, linear, two, Earth, second, SV, Jerks, several years, one, a1

File: S2212667814000380.txt

Extracted Keywords: ANSYS, 5950

File: S0021999113005718.txt

Extracted Keywords: 1, Monte Carlo, 2, continuum fluid

File: S0167931713004061.txt

Extracted Keywords: only one, PEGDMA, hydrogel, two, PAA, hydrogel scaffolds, first

File: S0370269304008792.txt

Extracted Keywords: 1.528A, the meson, 0.4

File: S0045782515001322.txt

Extracted Keywords: 1, 2 s, One, vessel, seconds, FE, around 30 mm, hours, FM

File: S1877750315300119.txt

Extracted Keywords: 0, 34, One, LR, Stieltjes, Section 3.1, 1]D., Table 2, Chaospy, the Askey-Wilson

File: S002002551630384X.txt

Extracted Keywords: Theron, 22, 39, 8, T4, Santamaria, 30, T3, 5.3, 24

File: S0009261413004612.txt

Extracted Keywords: only one, $H = -eB(S_1(z) + S_2(z)) + |e| S \rightarrow 2 \cdot I \rightarrow$, 3, t) fs(B, second, 8, t) B, RP, up to two, one

File: S0098300413002185.txt

Extracted Keywords: Gege, 1998, NOMAD, Southern Germany, seasonally, Bi], Heege, 2013b, season, the SeaWiFS Bio-Optical Archive and Storage System

File: S0963869514000863.txt

Extracted Keywords: 0% to 100%, 10, two, EM, steel, sensor, 7,8

File: S0370269304006756.txt

Extracted Keywords: QCD, Fermi, 5, zero

File: S221267161200162X.txt

Extracted Keywords: PID, linear, Hammerstein

File: S0885230816300043.txt

Extracted Keywords: MLLR, GMM, 0.8%, 0.9%, HMM, 0.3%, 2.9%

File: S0885230816301759.txt

Extracted Keywords: English, 96.6%, 84.1%

File: S0009261409006666.txt

Extracted Keywords: Q00, two, 50fs, the photon echo signals, three

File: S2212667814001476.txt

Extracted Keywords: Interclouds, DepSky, Cloud Computing, QoS, Cloud, Cloud of Clouds

File: S0370269304008706.txt

Extracted Keywords: 15, $^9\text{Be}(d, \text{Karmanov}, 9, ^{12}\text{C}(d, 4.5, 85, 0.8, 80, \text{T20}$

File: S221266781200007X.txt

Extracted Keywords: Crocodile Matchet Limited, Tema, Ghana, 6.74kW

File: S0003491615001839.txt

Extracted Keywords: Klein, Gordon, Statistical Mechanics, the van der Waals, 17-19,14,16, 4, Maxwell, Eq, first

File: S0370269304006082.txt

Extracted Keywords: 168, dressings, Standard, 11,13, Gepner, 29, one, Letter,

12, first

File: S0375960113004568.txt

Extracted Keywords: 10 and 12, two, M3, M4, M5, Fig

File: S2212671612000716.txt

Extracted Keywords: two, three

File: S1364815216302122.txt

Extracted Keywords: 1999, the Strength Pareto Evolutionary Algorithm, optimization, Firstly, SPEA, Thiele, Zitzler, Secondly, Pareto

File: S0377221716301357.txt

Extracted Keywords: one, MPSP

File: S0370269304008731.txt

Extracted Keywords: Majorana, lepton, Z3, neutrino

File: S2214657115000179.txt

Extracted Keywords: 1, Nickel

File: S0003491615001505.txt

Extracted Keywords: 2.4, Creutz, 32, Section 2.6, Pauli, continuum, Section

File: S0009261412012365.txt

Extracted Keywords: 1=35fs, C10H12+, norbornene, CPD, DCPD, two, 26, Matlab, 22,23, Levenberg

File: S0927025614006181.txt

Extracted Keywords: Verlet, 8, particle, The Discrete Element Method, half, one, Nottingham, al., Asmar

File: S0301010409001219.txt

Extracted Keywords: the recent years, fragmentation, 14,18,22

File: S2212667814001245.txt

Extracted Keywords: 980, Bayesian Network, Bayesian, Vietnamese, one

File: S0010938X15300512.txt

Extracted Keywords: 1, alumina, film, 8, pores, Al3, alumina films

File: S0301010415002189.txt

Extracted Keywords: 6, Martin, 10, DFT, JT, 8, 1978, SCF, 9, PES

File: S0957417416301786.txt

Extracted Keywords: 6, Firstly, 3, Huang & Wang, D-SDC, 2006, PLS, SW-SVR, Fig

File: S0301010415002256.txt

Extracted Keywords: 6, 13, decades, 5, Hamiltonian

File: S0168365913003295.txt

Extracted Keywords: one, only two, T-cells

File: S0022311515303640.txt

Extracted Keywords: 1,2, 3, bcc, bubble, 5, He, bubbles

File: S2212667814000070.txt

Extracted Keywords: Image Rotator, at least 30, Image Scaler, Color Space Converter, Image Resizer, Candidate Detector, Adaboost, smart TV, second, MCT Generator, Confidence Switch

File: S0370269304007634.txt

Extracted Keywords: 10, 11, Hooft, Einstein, Dirac, Polyakov, graviton, Montonen, Olive, Maxwell

File: S1071581916300854.txt

Extracted Keywords: 2014, Karrenbauer, Levine, 2015, Dunlop, 2012, Wiseman, 2013, Leiva et al., al.

File: S0038092X14004770.txt

Extracted Keywords: shading, 1, 2013, UK, Fig

File: S0370269304007695.txt

Extracted Keywords: hadron, FSI, the HERMES Collaboration, AGS, antiproton, 17, PYTHIA

File: S221266781300083X.txt

Extracted Keywords: one, Pareto, yield

File: S2212667814000690.txt

Extracted Keywords: Isochrysis galbana, 1, 24h, TISO, 18-20

File: S0888327016300048.txt

Extracted Keywords: two, EMD

File: S1570870516301822.txt

Extracted Keywords: 10, two, MWSN, 2, MANET

File: S0032386107010518.txt

Extracted Keywords: polymerization, 33, 3, Huisgen, polymers, ATRP, Opsteen, nucleophilic substitution, van Hest, 31

File: S0009261413011111.txt

Extracted Keywords: two, 4, 1-3, three

```
[ ]: from sklearn.feature_extraction.text import CountVectorizer, TfidfVectorizer
import json
#from rake_nltk import Rake
from sklearn.linear_model import LogisticRegression

texts = [open(os.path.join(text_dir, filename)).read() for filename in os.
    ↳listdir(text_dir)]
features = [extract_features(text) for text in texts]

# vectorizer = CountVectorizer()
vectorizer = TfidfVectorizer(max_df=0.8, min_df=0.2)

# X = vectorizer.fit_transform([json.dumps(feature) for feature in features])
X = vectorizer.fit_transform(texts)

model = LogisticRegression()
model.fit(X, extracted_keywords)

model.save("keyword_extraction_model.pkl")
```