

Design of Aerosol Literature Knowledge Service System*

Anonymous Author(s)[†]

ABSTRACT

This system is based on the website system of the historical literature database of aerosol field, which realizes the electronic, digital and online research of traditional review literature research. And use crowdsourcing and web-based data submission methods to effectively implement formatted submission of paper data to ensure database update and expansion. At the same time, through the thesaurus and history, enhance the user input experience and save data in time. The system also developed the corresponding online aerosol literature knowledge service system, and provided special services such as “document positioning, thesis topic classification, paper review, paper evaluation, paper journal recommendation, statistical analysis and visual display”, and introduced crowdsourcing submission. The model effectively provides relevant historical scholars with aerosol historical observation data services, laying a solid foundation for attracting more scholars to understand and participate in the construction of this system (data convergence).

CCS CONCEPTS

- Computer systems organization → Embedded systems; Redundancy; Robotics;
- Networks → Network reliability;

KEYWORDS

Aerosol; Database; Knowledge base; Document knowledge service system; Crowdsourcing; User experience

^{*}Produces the permission block, and copyright information

[†]The full version of the author's guide is available as acmart.pdf document

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

WOODSTOCK'97, July 2019, El Paso, Texas USA

© 2019 Association for Computing Machinery.

ACM ISBN 123-4567-24-567/08/06...\$15.00

https://doi.org/10.475/123_4

ACM Reference Format:

. 2019. Design of Aerosol Literature Knowledge Service System: Anonymous Author(s). In *Proceedings of ACM Woodstock conference (WOODSTOCK'97)*. ACM, New York, NY, USA, 6 pages. https://doi.org/10.475/123_4

1 INTRODUCTION

Aerosols play an important role in regulating climate change. In the context of global climate change, this regulation is global. To understand the impact of aerosols on global climate change, global aerosol observations are urgently needed.

At present, the observation data of aerosols are still scattered in a large number of documents and other types of storage media. For readers who need data, it is a cumbersome and unpleasant process to move between data platforms. Moreover, the traditional literature review is fixed in the form of papers. For the author, it is necessary to read a large amount of literature for statistical compilation, and every few years need to be updated to ensure the real-time statistical knowledge of such documents. The workload is large and cumbersome. For the reader, in order to quickly acquire such statistical knowledge, it is necessary to search for the review literature and judge their authority. The acquired knowledge is also limited by the region, year, etc. that the author wants to display.

In the face of a large number of intricate data, relying on traditional literature collection methods, it is difficult to extract aerosol data in massive documents quickly and efficiently, and the practice of reviewing for many years is far from satisfying the data iteration. In this article, we will present a solution to the problem by building an aerosol knowledge base. And on the basis of the establishment of the knowledge base, we have built the world's first comprehensive literature knowledge online website system dedicated to the aerosol field, giving full play to the role of literature data.

This article is divided into six parts to introduce this website system, the first section: related work; the second section: knowledge base design; the third section: knowledge service design; the fourth section: user experience; the fifth section: data expansion; Section: Conclusion.

107 2 RELATED WPRK

108 Our work intends to leverage the knowledge and services
 109 of the literature data, and we have built the world's first
 110 comprehensive online knowledge website system for the
 111 aerosol field. These technologies have attracted many efforts
 112 from different research fields. In this section, we will review
 113 the work in each area, as our work is inspired by the literature
 114 and draws lessons from previous methods.

116 Knowledge Base

118 Since 2002, R.Crow first proposed the concept of Institutional
 119 Repository (IR) to actively build institutional knowledge
 120 bases in universities and institutions libraries at home and
 121 abroad. It has been nearly 16 years, but the development
 122 of knowledge base is still trying. And groping, there is no
 123 ready-made road map. This is also a big problem that the
 124 system needs to solve in the design.

125 At present, the commonly used knowledge bases in China
 126 include Zhiwang, Wanfang, Baidu Academic, and the knowl-
 127 edge bases of major universities. There are Web of science
 128 and Nature in the world. However, most of them only pro-
 129 vide the functions of paper search and keyword indexing.
 130 Some domestic websites can provide services for checking
 131 papers. However, such data search methods are too broad,
 132 and they cannot quickly acquire a certain field in a certain
 133 period of time. Data trends within and statistical analysis
 134 results. This is exactly what we are trying to solve.

136 Knowledge Service Model

138 The current society is a service-based society, and the prod-
 139 uct production economy is gradually transformed into a
 140 service economy. The traditional innovation model based
 141 on technology is shifting from a user vision. We should also
 142 shift our product-centricity to users' center.

144 Crowdsourcing

146 The crowdsourcing method can recruit mass participants
 147 through the Internet to complete the work that the machine
 148 is difficult to accomplish. Since Howe proposed the con-
 149 cept of crowdsourcing in 2006, the research and applica-
 150 tion of crowdsourcing has developed rapidly. Due to the
 151 diverse backgrounds, low labor costs, and fast task comple-
 152 tion, crowdsourcing has developed rapidly and is widely used
 153 in image classification, manufacturing, film and television,
 154 human-computer interaction, and medicine.

155 In this article, we will also apply the crowdsourcing method,
 156 invite experts and scholars to fill out and submit the data of
 157 published articles, so as to expand the database.

160 3 IR DESIGN

161 The construction of IR requires good management, requires
 162 continuous planning, prioritization and coordination of dif-
 163 ferent stakeholders, including academic and academic re-
 164 searchers, libraries, institutional managers, publishers, and
 165 even students participating in research (doctors, masters)
 166 benefits.

167 The existence of IR must meet the needs of all parties
 168 involved. For this system, we plan, prioritize and coordinate
 169 different stakeholders, including experts in academic fields,
 170 teaching and research personnel in institutions, students par-
 171 ticipating in research (doctors, masters), and the interests of
 172 the masses, and analyze the interests. The difference between
 173 the reality and the needs of the party. See Table 1 for details.

174 The current literature data of this aerosol knowledge base
 175 covers 75 core aerosol-related journals, with more than 13,000
 176 aerosol-related literature data and more than 100,000 abstract
 177 data. The paper data covers China, South Korea, Japan, India,
 178 Africa, the United States and other countries and continents,
 179 as shown in Figure 1.

180 In China, the data covers Beijing, Chongqing, Xinjiang,
 181 Jiangsu Province, Anhui Province, Taiwan Province, Guang-
 182 dong Province, Hubei Province, Guizhou Province, Gansu
 183 Province, and Shaanxi Province, as shown in Figure 2.

186 4 KNOWLEDGE SERVICE

187 We regard knowledge service as a system problem. Starting
 188 from the user's needs mining, through the form of ques-
 189 tionnaires and interviews, we collect and determine the re-
 190 quirements, systematically apply the theory and method of
 191 design to create and plan services: system for regions, wave-
 192 lengths , the year and other conditions for statistical analysis
 193 of the data in the database, and visual display, hope that users
 194 can quickly and intuitively obtain statistical knowledge of
 195 aerosol data. Services form a process and have value to end
 196 users, resulting in high quality services that enhance the
 197 user experience.

198 Different services have different meanings at different
 199 times, so service products are often personalized. In this re-
 200 gard, this paper designs adaptive service functions for the
 201 knowledge service system based on the exclusive character-
 202 istics of the aerosol literature. Therefore, after we get Table 3,
 203 we made relevant user surveys for the personnel objects in
 204 Table 3 and their interests, in order to determine the service
 205 function points.

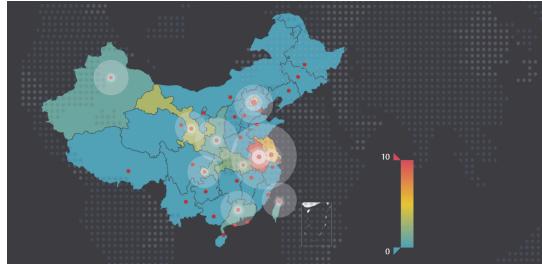
208 User Questionnaire

209 We set up an online questionnaire to collect public informa-
 210 tion. A total of 20 questionnaires (currently) were collected,

107	2 RELATED WPRK	160
108		161
109		162
110		163
111		164
112		165
113		166
114		167
115		168
116		169
117	Knowledge Base	170
118		171
119		172
120		173
121		174
122		175
123		176
124		177
125		178
126		179
127		180
128		181
129		182
130		183
131		184
132		185
133		186
134		187
135		188
136		189
137	Knowledge Service Model	190
138		191
139		192
140		193
141		194
142		195
143		196
144		197
145	Crowdsourcing	198
146		199
147		200
148		201
149		202
150		203
151		204
152		205
153		206
154		207
155		208
156		209
157		210
158		211

213
214
215
216
217
218
219
220
221
222
223

(a) World Map

224
225
226

(b) China Map

227

Figure 1: Subscale.

228

Table 1: Tentative Function Points Questionnaire

229

Functions	Proportion
Keyword Query	xx%
Topic Classification	xx%
Paper Review	xx%
Paper Evaluation	xx%
Paper Template	xx%

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

including 3 in the aerosol professional field and 17 in the non-aerosol professional field (both students). The questionnaire was in the form of a question, including but not limited to:

(1) What is the method for determining the direction of the paper before the paper is written?

(2) In what form do you want to get data in the paper?

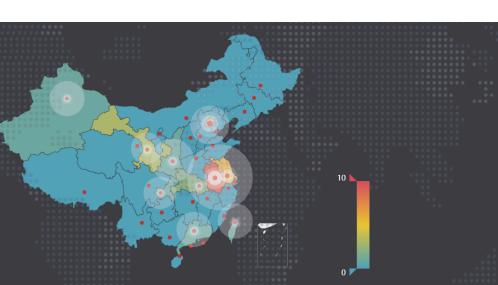
(3) In what form do you want to judge your paper after the paper is written?

For this survey, we got a lot of important responses and suggestions. In the process of sorting out the answers, we made clear the determination of the system function. The tentative function points are as follows in Table 2:

User Interview

In order to further deepen the user's needs and determine the system functions, we adopted the focus group interview method. We interviewed 100 volunteers, including 50 in the aerosol field (professionals include teachers and students) and 50 non-aerosol professionals (occupations include teachers, students, and other corporate workers).

Face-to-face user interviews were conducted with a group of 10 people (the same professional). We will first set up different aerosol literature knowledge related project questions based on volunteer career information, record the problem solving needs through the discussion between volunteers, and let the volunteers click on the importance of the functions we selected before. The scale is shown in Figure 3:



(b) China Map

Figure 1: Subscale.

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

48



Figure 3: Main page.

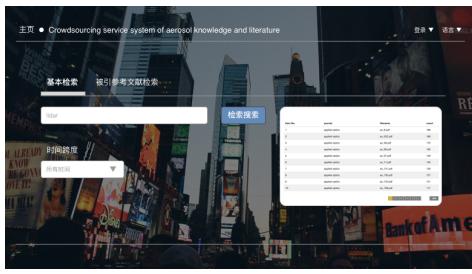


Figure 4: Information search.

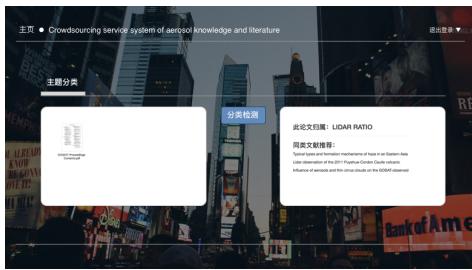


Figure 5: Subject classification.

descending order, click to view the documents, and automatically locate the fields of the searched keywords. As shown in Figure 4.

Subject Classification. Users can upload documents, the system will accurately locate the document level and category, and inform the document subject classification. According to the similarity of the topic, the relevant literature is recommended for the user. The user can click to view the recommended documents. As shown in Figure 5.

Paper Check. Users can upload his own document and the system will upload the PDF document as the target file, perform data preprocessing such as word segmentation and deactivation, and convert it into a txt file, compare the similarity with other documents in the library, and feedback the user and upload the document. Similar published literature, users can click to view the literature, the system will



Figure 6: Paper check.

automatically locate similar fields in the literature, which is convenient for users to view. As shown in Figure 6.

Data Crowdsourcing. See the next section for details.

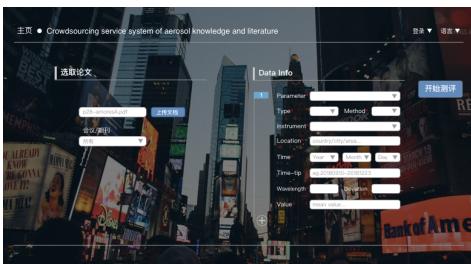
Paper Evaluation. When users upload a manuscript written by themselves, the system will use the uploaded PDF document or the submitted paper information as the target files, and compare them with the published papers or the more influential papers in the library. Finally, feedback the level of the user's target paper and recommend the corresponding excellent papers. As shown in Figure 7.

5 CROWDSOURCING AND DATA EXPANSION

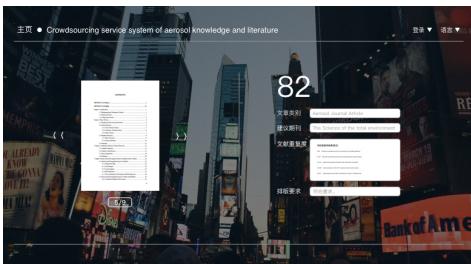
As system users continue to grow and data continues to emerge, data update iterations are a major issue that needs to be addressed. In the subsequent data collection and update of the knowledge service system in this document, we adopted a crowdsourcing method and proposed a new formatted data submission method to combine crowdsourcing submission data with data formatting storage and data analysis to achieve perfect knowledge. The role of the library and feedback user data information to ensure the update and expansion of the database.

Parameter Setting

After discussing with the graduate students of the Optoelectronics College Collaborative Laboratory, we designed a questionnaire experiment. The subjects were 30 graduate students in the School of Optoelectronics. The questionnaire was divided into five parts: literature publication information, aerosol basic information, optical parameter information, spatio-temporal information, and particle microphysical characteristics. The subjects were required to give these five parts separately. I think that an aerosol data related literature may contain important attributes. We compiled and merged the attributes of the 30 questionnaires collected. The results of the questionnaire included preliminary 28 attributes that would cover the important information needed. To increase the scalability, we added a "Remarks" attribute



(a) Data submission



(b) Evaluation result

Figure 7: (a) In the [Data info] list, the user needs to fill the list with aerosol parameter type, Method, Instrument, Location, Time, Time-tip (such as a certain time period, a certain season, etc.), Wavelength, Deviation, Value (parameter other Values, such as Mean_value, etc.). In addition to the Time-tip and Deviation options, the rest are required information items. (b) On the right side of the interface, the system will give a rating item, including Overall score; paper category; recommended journal: give the name of the journal that best meets the submission; document repeatability: list the top five documents with the highest similarity to the uploaded paper in the whole library, sorted in descending order of similarity; : Compare the paper format with the existing paper format template in the database to determine whether it meets the requirements.

to each category. For the recording of other possible additional information, the total attribute is 33 as a paradigm preliminary attribute option, as shown in Figure 9.

In the follow-up, we further interviewed experts and scholars in the field of aerosols, taking into account the importance of the data and the clarity of the records. Finally, the attributes included in the paradigm were determined to be 17 and the four categories except the published information were combined. Into the data, the attributes contained in the paradigm are divided into article attributes and data attributes in the large class. As shown in Tables 4 and 5.

Web Design

The system uses the eight keywords of lidar ratio, depolarization ratio, backscatter coefficient, extinction coefficient, optical

Table 3: Article Attribute.

Article Attribute	Content
Journal	75 journal options
DOI	Unique Number of article
Paper Name	Full name of the document
Author	Corresponding author
Unit	Company or school
Tool	Observation, simulation

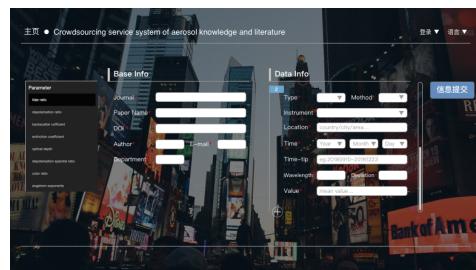


Figure 8: Website navigation.

depth, depolarization spectral ratio, color ratio, angstrom exponents to guide the document information from Journal, DOI, Type, Method and so on. Extract and fill in the submission database. The specific web page information is shown in Figure 8.

In terms of user experience, in order to save user data filling time, we set the thesaurus function, set the drop-down options in multiple options such as Journal, Instrument, Type, and province cities, users can click to complete. In terms of data security, the system will automatically save the user to fill in the data to avoid data loss caused by webpage error shutdown, network disconnection and other reasons. The specifics of the user interaction experience will be covered in Section 6.

Crowdsourcing task organization method

We submit data through invitations and autonomy.

By sending an email to invite experts and scholars in the aerosol field, journals and journals to submit the data of the published papers on the crowdsourcing submission page of the system, this will help improve the accuracy of the system's "paper evaluation" function. According to the format paradigm of excellent papers, we can judge the grades and advantages and disadvantages of the evaluation documents submitted by users.

Student users can get convenience from the system. After publishing the paper, we encourage and very welcome users to submit data independently and help the system to expand data.

Table 4: Data Attribute.

Data Attribute	Content	
Key Physics Parameters	Lidar ratio, depolarization ratio, backscatter coefficient, extinction coefficient, optical depth, depolarization spectral ratio, color ratio, angstrom exponents	584
Aerosol Type	Dust, smoke, clean continental, polluted continental, polluted dust, clean marine	585
Method	Observation, retrieval	586
Laser Wavelength	355, 440, 532, 645, 780, 870, 1020, 1540	587
Address	Country, province, city	588
Address Note	Such as a foreign country, the country specific to the school, latitude and longitude and other information	589
Time	Year, month, day	590
Time Note	Time interval	591
Mean	Optical parameter average, decimal form	592
Standard deviation	Optical parameter standard deviation, decimal form	593
Value note	Data Interval	594

Automatic system update

In addition to the background staff uploading the document data, the background of the system will automatically detect the recent literature update of each data platform, the crawler automatically downloads the document information and the abstract, promptly reminds the background personnel to update the data.

6 USER STUDY

We conducted an experimental test of the experience of the website. We invite volunteers who have participated in the questionnaires and interviews, as well as volunteers who have not participated in the activities related to the system, and provide us with feedback.

System usability and usability test

We convened the subjects who had previously interviewed to complete the system's usability and usability tests. The experimenter will experience the various functions of the system by looking up the paper information. Since our system is aimed at the field of photoelectrosols, our subjects were also professionals in the field of interviews.

Similarly, we use the subscale to let users rate the system's ease of use and usability, and make recommendations for system layout. The system usage is shown in Table 6.

System time-saving test

...

7 CONCLUSIONS

This paper proposes a new method of data collection, storage, collation, analysis and presentation of literature, and

Table 5: Article Attribute.

Article Attribute	Content	
Journal	75 journal options	604
DOI	Unique Number of article	605
Paper Name	Full name of the document	606
Author	Corresponding author	607
Unit	Company or school	608
Tool	Observation, simulation	609

develops a paradigm analysis and crowdsourced aerosol literature knowledge service system, which provides an effective means for collecting aerosol historical observation data and provides users with A variety of aerosol literature knowledge service functions.

At present, the system functions need to be improved. The system interaction and user experience need to be considered from the user's point of view. Through more tests, the system efficiency is improved.

ACKNOWLEDGMENTS

...

REFERENCES