

برنام خدا



پژوهشگاه علوم و فناوری اطلاعات ایران  
(ایزدانک)

## دشواری‌های سازماندهی اطلاعات



کلیدواژه‌ها



نام‌های جغرافیایی



نام پدیدآوران



کوتاه‌نوشته‌ها و سرنام‌ها



نام گیاهان و جانوران

## مقدمه

سازماندهی اطلاعات از شمار فعالیت‌های کلیدی پژوهشگاه علوم و فناوری اطلاعات برای پردازش پایان‌نامه‌ها و رساله‌های دانشجویی (پارسا) و دیگر انتشارات علمی است. با انبوه انتشارات و اطلاعاتی که امروزه پدید می‌آیند، سازماندهی آنها بیش از پیش نیاز است تا دسترسی جویندگان این اطلاعات آسان‌تر و با شتاب بیشتری فراهم شود. بدون سازماندهی، یافتن اطلاعات درست دشوار است، پیشرفت‌های علمی کند می‌شود، و تحلیل اطلاعات شدنی نیست. از این رو، کارشناسان و متخصصان، همه اطلاعاتی را که ایراندک گردآوری می‌کند به شکلی کارآمد سازماندهی می‌کنند. سازماندهی اطلاعات دربردارنده کنترل درستی اطلاعات، نمایه‌سازی، چکیده‌نویسی، و کارهایی از این دست است. ولی سازماندهی همیشه ساده نیست و گاه با چالش‌هایی همراه است. گوناگونی در نام گیاهان و جانوران، یک‌دست نبودن کوتاه‌نوشت‌ها و سرنام‌ها، تفاوت در نگارش نام پدیدآوران، گوناگونی در نام‌های جغرافیایی، و یک‌دست نبودن نگارش کلیدواژه‌های هم‌معنا در شمار کلیدی‌ترین چالش‌هایی هستند که کارشناسان سازماندهی اطلاعات ایراندک با آنها روبه‌رو می‌شوند. گزارش پیش‌رو دربردارنده توضیح این دشواری‌ها و نمونه‌هایی در این زمینه است.

## ۱. نام گیاهان و جانوران

نام گیاهان و جانوران از چالش‌های کلیدی سازماندهی اطلاعات در حوزه‌هایی است که با این نام‌ها سروکار دارند. از آنجا گیاهان و جانوران نام علمی دارند و اغلب میان نام علمی و نامی که عامه مردم روی آنها می‌گذارند تفاوت هست، پدیدآوران به شکل‌های گوناگون این نام‌ها را در آثار خود می‌آورند. ترجمه از زبان‌های دیگر به زبان فارسی نیز بر این دشواری می‌افزاید. از سوی دیگر، در منطقه‌های جغرافیایی گوناگون نیز شاید یک گیاه یا جانور نام متفاوتی داشته باشد. اصطلاح نامه‌های «کب» و «آگرووک» از نام علمی گیاهان و جانوران بهره برده‌اند ولی در اصطلاحنامه کشاورزی (فارسی) نام‌ها آوانویسی شده‌اند. جویندگان اطلاعات نیز با نام‌های گوناگون جست‌وجو می‌کنند. نمونه‌های زیر گویای دشواری سازماندهی این نام‌ها هستند.

## نمونه یک: چند نام برای یک گیاه



نام گیاه: لبدیسی خزنده (*Ajuga reptans*)  
نام‌های دیگر فارسی: لبدیسی رونده، لبدیسی گسترده، فرش منجوقی  
نام‌های دیگر انگلیسی: carpetweed, carpet bugle, blue bugle, bugleherb, bugleweed, bugleweed

## نمونه دو: چند نام برای یک گیاه



نام گیاه: آمیگدالوس اسکوپاریا (*Amygdalus scoparia*)  
نام‌های دیگر فارسی: بادام وحشی، بادام کوهی، ارژن (درخت)  
نام‌های دیگر انگلیسی: Mountain almond, wild almond, Prunus scoparia

## نمونه سه: چند نام برای یک جانور



نام جانور: بزماهی زرد جامه (*Upeneus sulphureus*)  
نام‌های دیگر انگلیسی: Sulphur Goatfish, Beach goatfish, Goatfish, Sunrise goatfish, Yellow goatfish  
نام فرانسوی: Rouget-souris aurore  
نام اسپانیایی: Salmonete aurora

**نمونه چهار: یک نام برای چند جانور**



جانور (شکل بالا): کفشک زبان گاوی (Speckled tonguesole, *Cynoglossus puncticeps*)  
 جانور (شکل پایین): کفشک زبان گاوی (Cynoglossus bilineatus, Fourlined tonguesole)  
 توضیح: برخی از پدیدآوران فارسی زبان در آثارشان برای دو شکل روبه‌رو از «کفشک زبان گاوی» بهره‌برداری می‌کنند، ولی شکل بالا «کفشک زبان گاوی لکه‌دار» و شکل پایین «کفشک زبان گاوی چهارخطه» نام‌گذاری شده است.

**نمونه پنج: چند نام برای یک درخت**



نام درخت: شیردار (*Acer Cappadocicum*)  
 نام‌های دیگر: شیردار (*Acer Cappadocicum*)، شیردار (گیاه) (*Coliseum Maple*)، آسر کاپادوسیوم (*Acer Cappadocicum*)، شیرپلت (*Acer Cappadocicum*)، کیکم، اج، بز برگ، و ...  
 توضیح: در منطقه‌های جغرافیایی گوناگون نیز با نام‌های متفاوت این درخت را می‌شناسند: رودبار، آستارا، طالش و درفک (ککم، کیکم)، گسلان و کلارستاق (پلت)، سیاه بیشه و پل زنگوله (شیرپلت)، لاهیجان، رودسر، و شهسوار (عاج، آج، و اج)، اسکوله گیلان (فاقاج)، گرگان (بزبرک) و میاندره (بزوالک). جهاد کشاورزی نیز رویه روشنی برای نام‌گذاری گیاهان ندارد.

**نمونه شش: بهره‌برداری کاربران از نام‌های گوناگون برای جست‌وجوی یک گیاه**



نام گیاه: سیاهدانه (*Negella*)  
 نام‌های دیگر انگلیسی: Fennel-flower, Nigella, Golden thread, Devil in a brush, Garden fennel-flower, Black cumin, Nutmeg flower, Roman coriander, black caraway, black sesame, black seed  
 توضیح: سیاهدانه در ایران هشت گونه علفی دارد: سیاهدانه ایرانی، سیاهدانه ناجوربرگ، سیاهدانه مزرعه روی، سیاهدانه دمشقی، سیاهدانه هرز، سیاهدانه سیاهدانه‌ای، سیاهدانه سوری، سیاهدانه؛ شونیز. کاربرد به فارسی به دنبال نام علمی آن سیاهدانه، یعنی «نیژلا ساتیو» نیست، او به دنبال سیاهدانه و گونه‌های آن است. اگر داده‌ها به شکل کامل در اصطلاحنامه وارد شوند می‌توان میان نام‌های انگلیسی و فارسی پیوند ایجاد کرد.

**۲. کوتاه‌نوشت‌ها و سرنام‌ها**

کوتاه‌نوشت‌ها و سرنام‌ها از دیگر چالش‌های سازماندهی اطلاعات هستند. گاهی برای یک عبارت چند کوتاه‌نوشت یا سرنام هست، گاهی از آوانویسی کوتاه‌نوشت‌ها و سرنام‌های انگلیسی در فارسی بهره‌برداری می‌شود، و گاهی هم میان کوتاه‌نوشت‌ها و سرنام‌های فارسی و انگلیسی تفاوت هست. نمونه‌های زیر گویای دشواری سازماندهی کوتاه‌نوشت‌ها و سرنام‌ها هستند.

**نمونه یک: یک سرنام برای چند عبارت**

- جنبش مقاومت اسلامی (حماس) به کاربرد: حماس
- حرکت مقاومت اسلامی (حماس) به کاربرد: حماس
- جنبش مقاومت اسلامی به کاربرد: حماس
- حماس (جنبش مقاومت اسلامی) به کاربرد: حماس

- Harakat al-Muqawamah al-Islamiyah (Hamas) USE Hamas
- Islamic Resistance Movementcarpet bugleweed USE Hamas

**نمونه دو: یکدست نبودن عبارتهای کوتاه‌نویسی شده**

ایلوئیس (نرم افزار)

 ILWIS (software) سامانه یکپارچه اطلاعات زمین و آب (ایلوئیس) Integrated Land and Water Information System (ILWIS)**نمونه سه: کاربردی نبودن اعراب‌گذاری در جست‌وجو**

معتادان جنسی گمنام: sex addicts anonymous

 مُجگم به کاربردید: معتادان جنسی گمنام معتادان جنسی گمنام (مجگم): sex addicts anonymous (SAA)**نمونه چهار: همانند نبودن سرنام فارسی و انگلیسی**

مشاوره و آزمایش داوطلبانه (ماد): Voluntary Counselling and Testing (VCT)

 ماد به کاربردید: مشاوره و آزمایش داوطلبانه مشاوره و آزمایش داوطلبانه (ماد) به کاربردید: مشاوره و آزمایش داوطلبانه مشاوره و آزمایش داوطلبانه (وی.سی.تی.): Voluntary Counselling and Testing (VCT)**نمونه پنج: چند سرنام فارسی برای یک سرنام انگلیسی**

H9N2 Subtype Influenza A Virus

برابرنهاده‌های پیشنهادی فارسی:

 اچ‌ان۲ نوع دوم ویروس آنفلوآنزا نوع آ اچ‌ان۲.۹ نوع دوم ویروس آنفلوآنزا نوع آ اچ‌ان۲-۹ ویروس آنفلوآنزای نوع آ. اچ‌ان۲ نوع دوم ویروس آنفلوآنزا نوع آ.**نمونه شش: سرنام‌ها و آوانویسی‌های گوناگون** Open Computing Language (OpenCL) OpenCL (Open Computing Language)

برابرنهاده‌های فارسی:

 اپن سی ال اپن سل اپن.سی.ال. اپن.سی.ال. اپن.سی.ال اپن سی ال (رابط برنامه نویسی) زبان برنامه‌نویسی باز (اپن سی ال)**نمونه هفت: تفاوت در سرنام‌های فارسی و انگلیسی**

داعش:

 Islamic State of Iraq and the Levant (ISIS) Islamic State (IS) Islamic State of Iraq and Syria Islamic State of Iraq and al-Sham (ISIS)



□ داعش (دولت اسلامی عراق و شام): Daesh (Islamic State of Iraq and the Levant)

□ دولت اسلامی عراق و شام (داعش): ISIS (Islamic State of Iraq and the Levant)

□ داعش: Daesh

□ داعش: IS (Organization)

### ۳. نام پدیدآوران

نام پدیدآوران همیشه از چالش‌های سازماندهی اطلاعات بوده است. نمایه‌نامه‌های بزرگی همانند «وب آو ساینس» و «اسکوپوس» نیز هنوز به شکل کامل نتوانسته‌اند پاسخ کارآمدی به این چالش دهند. اگرچه با دادن یک شماره یکتا به هر پدیدآور کوشیده‌اند تا جایی که شدنی است یک‌دستی در نام پدیدآوران ایجاد کنند. پایگاه‌های اطلاعاتی ایرانداک دربردارنده بیش از نهصد هزار نام پدیدآوران هستند. کارشناسان ایرانداک نیز با دشواری سازماندهی نام پدیدآوران دست‌وپنجه نرم می‌کنند. نمونه‌های زیر گویای این دشواری هستند.

#### نمونه یک: شکل‌های گوناگون نگارش نام یک پدیدآور

□ ولی الله	بابازاد	□ ناصر	علیقلی زاده فیروز جایی	□ مهرانگیز	علی نژاد محمد آباد
□ ولی الله	بابازاد	□ ن.	علیقلی زاده فیروز جایی	□ مهرانگیز	علی نژاد محمد آباد
□ ولی الله	بابازاد	□ -	علیقلی زاده فیروز جاهی	□ مهرانگیز	علی نژاد محمد آبادی
□ ولی الله	بابا زاد	□ ناصر	علیقلی زاده فیروز جاهی	□ مهرانگیز	علی نژاد محمد آبادی
□ ولی اله	بابا زاد	□ ناصر	علی قلی زاده فیروز جاهی	□ مهرانگیز	علی نژاد محمد آبادی
□ ولی اله	بابازاد	□ ناصر	علی قلیزاده فیروز جاهی	□ مهرانگیز	علی نژاد محمد آبادی
□ ولی ...	بابازاد	□ ناصر	علیقلی زاده فیروز جاهی	□ مهرانگیز	علی نژاد محمد آبادی
□ ولی ..	بابا زاد	□ ناصر	علیقلی زاده فیروز جایی		
		□ ناصر	علیقلی زاده فیروز جایی		
		□ ناصر	علیقلی زاده فیروز جایی		

#### نمونه دو: تفاوت در نام خانوادگی یک پدیدآور

- تیمور لنگ
- تیمور گورکانی
- تیمور گورکان
- امیر تیمور
- امیر تیمور گورکانی
- تیمور
- تامرلن

### ۴. نام‌های جغرافیایی

چالشی که درباره نام پدیدآوران در سازماندهی اطلاعات هست، برای نام‌های جغرافیایی نیز پیش می‌آید. در چند سال گذشته سازمان نقشه‌برداری کشور شیوه‌نامه‌ای برای آوانویسی نام‌های جغرافیایی تهیه کرده است که شاید در آسان کردن برخی از دشواری‌ها کارآمد باشد. نمونه‌های زیر گویای این دشواری هستند.

**نمونه یک: چند نام برای یک منطقه جغرافیایی**

Buin va Mian Dasht □ بوئین و میاندشت	□ مشگین شهر Meshgin Shahr	Islamabad □ اسلام آباد	□ تپه قبرستان سر ( فولاد محله ) □ Tappeh Ghabrestan sar (Fooladmahalleh)
Booin & Miyandasht □ بوئین و میاندشت	□ مشگین شهر Meshginsahar	Eslam Ababd □ اسلام آباد	□ Ghabrestan sar Tappeh (Foolad Mahalleh)
Bouin va Mian Dasht □ بوئین و میاندشت	□ مشگین شهر Mešginšahr	Islam Abad □ اسلام آباد	□ Qabrestansar Tappeh (Foolad Mahalleh)
Bouin and Miandasht □ بوئین و میاندشت	□ مشگین شهر Meshginsahar	Eslamabad □ اسلام آباد	□ Ghabrestan sar Hill (Fooladmahalleh)
Booein va Miandasht □ بوئین و میاندشت			□ Qabrestansar Hill (Foulamahalleh)
Bu'inmiyāndašt □ بوئین و میاندشت			□ Tappeh Ghabrestansar (Fooladmahalleh)

**نمونه دو: تغییر در ارجاعها به منطقه‌های جغرافیایی**

□ کرمانشاه به کار برید: باختران	□ خسرو شاه به کار برید: خسروشهر
□ باختران به کار برید: کرمانشاه	□ خسروشهر به کار برید: خسروشاه

**۵. کلیدواژه‌ها**

نمایه‌سازی اطلاعات بیشتر با کلیدواژه‌هایی شکل می‌گیرد که پدیدآور یا نمایه‌ساز به انتشارات می‌دهد. اگر این کلیدواژه‌ها درست و یک‌دست گزینش شوند سازماندهی اطلاعات آسان‌تر و بازیابی آن کارآمدتر است. ولی همیشه اینگونه نیست و در موارد بسیاری این کلیدواژه‌ها درست و یک‌دست گزینش نمی‌شوند. نمونه‌هایی از دشواری کار با کلیدواژه‌ها در ادامه آمده‌اند.

**نمونه یک: چند کلیدواژه برای یک مفهوم**

□ فضای مجازی	□ توزیع دم - سنگین	□ گاستروآنتریت	□ تازگی، تکرار و ارزش پولی (آ.ر.اف.ام.)
□ رایاسپهر (پیشنهاد فرهنگستان)	□ توزیع دنباله - سنگین	□ التهاب معده روده	□ آ.ر.اف.ام. (ارزش مشتری)
□ فضای سایبری	□ Heavy-tail distribution	□ التهاب معده روده‌ای	□ Recency, Frequency and Monetary value (RFM)
□ Cyberspace		□ التهاب معده‌ای روده‌ای	
		□ Gastroenteritis	
		□ Gastroenteritis (G.E)	

**نمونه دو: تفاوت در آوانویسی کلیدواژه‌های انگلیسی**

□ تحلیل تکافت ( فرهنگستان زبان و ادب فارسی ) SWOT analysis	□ روش آموس
□ تحلیل توانایی - کاستی - فرصت - تهدید (سوات)	□ روش ایموس
□ تحلیل سوات SWOT	□ روس ای موس
□ تحلیل اس دبلیو اُ تی	□ روش اموس
□ روش راهبردی SWOT	□ روش روزآمدسازی خودکار چند رسانه ایی توسعه جایگشتی
□ Strengths, Weaknesses, Opportunities, and Threats analysis (SWOT)	□ به روزرسانی چند زمانه خودکار به روش بسط شاخص انعکاس طیفی
	□ AMUS method
	□ Automated Multi-temporal Updating by Signal Extension