



Cómo captar estudiantes a través de una buena arquitectura del conocimiento para tu universidad

Un experimento de 90 días.

David Martínez Mata

Máster en FLE y Ciencias del Lenguaje. Sorbonne Université, París.

Máster Hispanofrancés en Lengua Francesa Aplicada. Universidad Complutense de Madrid.



09 de febrero de 2019



**¿Qué sabe un aspirante
sobre tu universidad?**



**¿Qué sabe la comunidad
sobre tu universidad?**



**¿Qué sabe YouTube
sobre tu universidad?**



**¿Qué sabe Google
sobre tu universidad?**



**¿Qué sabe Chat GPT
sobre tu universidad?**



No saben nada o muy poco.



Relato 1

125.218 URLs: el reino del caos



**El dominio www.uniandes.edu.co/ fue
registrado en enero de 1996**

Las webs de las universidades tienen casi 30 años

28

años en promedio

14 universidades colombianas.
Dominios registrados entre 1996 y
2000.

Los Andes
uniandes.edu.co · 01-01-1996 **30** años

La Sabana
unisabana.edu.co · 14-04-1998 **28** años

Tecnológica de Pereira
utp.edu.co · 28-04-1998 **28** años

del Norte
uninorte.edu.co · 16-07-1998 **28** años

del Valle
univalle.edu.co · 16-07-1998 **28** años

EAFIT
eafit.edu.co · 17-07-1998 **28** años

Pontificia Bolivariana
upb.edu.co · 21-07-1998 **28** años

Nacional
unal.edu.co · 23-07-1998 **28** años

Externado
uexternado.edu.co · 26-08-1998 **28** años

del Rosario
urosario.edu.co · 02-09-1998 **28** años

Militar Nueva Granada
umng.edu.co · 01-10-1998 **28** años

Icesi
icesi.edu.co · 13-11-1998 **27** años

Javeriana
javeriana.edu.co · 06-07-1999 **27** años

de Antioquia
udea.edu.co · 04-02-2000 **26** años

Casi tres décadas publicando contenido

¿Qué hay en el universo de un sitio web?

U. de las Ciencias Básicas

106.230

URLs rastreadas

73.387 URLs HTML

U. de la Salud

11.895

URLs rastreadas

5.550 URLs HTML

U. del Derecho

7.093

URLs rastreadas

2.095 URLs HTML

La arquitectura de contenidos empieza por saber qué existe, dónde está y cómo se llama.

¿Qué tipo de recurso encontramos en cada web?

U. de las Ciencias Básicas

69 %

páginas HTML

7 % imágenes · 24 % otros

U. de la Salud

49 %

imágenes

48 % páginas · 3 % PDF

U. del Derecho

53 %

PDF

32 % páginas · 13 % imágenes

Cada web acumula un recurso distinto: páginas, imágenes o documentos.

Pero muchas páginas con errores

U. de las Ciencias Básicas

36 %

son URLs en bucle

enlaces mal contruidos que generan rutas
infinitas

U. de la Salud

13,5 %

dan error 404

sobre todo complementos de programa

U. del Derecho

35 %

dan error 404

607 son enlaces a PDF que ya no existen

Antes de ordenar, hay que saber cuánto de lo publicado ya no sirve.

¿Qué parte de las URLs es oferta académica?

U. de las Ciencias Básicas

0,2 %

de todas las URLs son programas

249 programas en 106.230 URLs

U. de la Salud

2,0 %

de todas las URLs son programas

236 programas en 11.895 URLs

U. del Derecho

4,6 %

de todas las URLs son programas

329 programas en 7.093 URLs

Oferta académica: una URL principal por programa (pregrado, posgrado y educación continua).

¿Cuántas URLs están relacionadas con cada programa?

U. de las Ciencias Básicas

4,8

URLs de apoyo por programa

variantes, complementos, listados y documentos

U. de la Salud

7,8

URLs de apoyo por programa

1.016 complementos: plan de estudios, docentes, perfiles

U. del Derecho

1,3

URLs de apoyo por programa

293 PDF de programa y casi ninguna página alrededor

Documentos que podrían servir si supiéramos qué son

U. de las Ciencias Básicas

79 %

de los PDF no tiene título interno

`20260918092700_0.pdf`

U. de la Salud

56 %

de los PDF no tiene título interno

título interno: «Document»

U. del Derecho

83 %

de los PDF no tiene título interno

`auto-admite-11001-33-35-010-2
025-00036-00.pdf`

Si el documento no dice qué es, nadie lo encuentra para reutilizarlo.

¿Con qué criterio está guardado hoy?

U. de las Ciencias Básicas

99 %

de los PDF bajo una ruta técnica

`/showimagen/v/Comunicaciones/`

U. de la Salud

92 %

de los PDF en carpetas por mes

`/sites/default/files/2026-07/`

U. del Derecho

100 %

de los PDF en carpetas por mes

`/wp-content/uploads/2025/09/`

0 carpetas por programa, facultad o tema en las tres.

¿Las imágenes dicen lo que muestran?

U. de las Ciencias Básicas

78 %

de las imágenes tiene nombre
genérico

`20230829122037_0.jpg`

U. de la Salud

18 %

de las imágenes tiene nombre
genérico

`WhatsApp Image 2026-05-07
at 7.06.38 AM.jpeg`

U. del Derecho

37 %

de las imágenes tiene nombre
genérico

`df93c16368e5858b568bff9c
c6b2d494.png`

El nombre no dice qué hay en la imagen.

¿De qué tratan los PDF que sí se identifican?

U. de las Ciencias Básicas

9 %

institucional

informes, certificados, acreditación

U. de la Salud

25 %

institucional

más 20 % normativa y 9 % costos

U. del Derecho

52 %

judicial

tutelas, autos, demandas: 71 % de todas las páginas

Pocos PDF son académicos: 2 %, 5 % y 8 %.

Para empezar una estrategia, tenemos tres inconvenientes:

1 Clasificación

Existe un desconocimiento de dónde viven los recursos.

2 Nombramiento

Los textos carecen de un título o descripción propia.

3 Rastreo

Entre el 13,5 y el 36% de las URLs están rotas

Si ni la universidad encuentra su contenido, tampoco un aspirante ni un agente de IA.



**Si la IA sabe muy poco sobre mi
estrategia de contenidos,
¿cómo debo crear el contenido para
que me entienda mejor?**



Relato 2

La página que nadie nunca quiso escribir

Creación de un protocolo de explicitación semántica (PES)

La hipótesis del proyecto

¿Explicitar las relaciones léxico-semánticas de un texto aumenta la probabilidad de que un sistema de IA lo recupere y lo cite?



Variable independiente: el Protocolo de Explicitación Semántica (PES) aplicado o no a la redacción.

Dos sitios similares: mismo servidor, mismo día, sin enlaces



Corpus A - Optimizado con PES

apprendreunelangue.com

Carpeta: corpus_pes/



Corpus B - Sin optimización sin PES

commentdevenirbilingue.com

Carpeta: corpus_standard/

Qué se mantiene igual

- Fecha de creación y servidor.
- Montado en WordPress y con la misma jerarquía de URLs.
- Cero backlinks.
- Mismas 10 escuelas por ciudad.
- Mismo modelo de redacción (GPT-4o).
- Mismas fichas, reseñas y enlaces a Maps.

¿Qué cambia en el proyecto? Las instrucciones de redacción (¿cumple o no con el PES), los encabezados H1 y H2, los metadatos, los cuales están estructurados de manera distinta para que se diferencien, de forma que Google no los trate como contenido duplicado.

Recoger las páginas que Google ya posiciona

Para cada ciudad se consulta la SERP orgánica de Google Francia y se guardan las URLs orgánicas sin duplicados. Es el corpus de referencia del que sale el vocabulario.

```
POST /v3/serp/google/organic/live/advanced
[ {
  "keyword": "meilleures écoles anglais Rennes",
  "location_code": 2250,    # Francia
  "language_code": "fr",
  "depth": 30              # 20 en expansión
}]
```

¿De dónde proviene la data? DataForSEO

Bucle de expansión

- 1 Ronda 1: keyword semilla, la cual arroja un total de 30 resultados.
- 2 Paso 2: extrae términos y las convierte en consultas nuevas para la segunda ronda de consulta en las SERPs.
- 3 Ronda 2+: cada consulta nueva, lo que arroja 20 resultados.
- 4 Las URLs ya vistas se descartan.

Rennes: 90 URLs, incluyendo las primeras 20 URLs más las extraídas por la expansión de keywords.

lexical_config.json: los términos semánticos por ciudad

```
{
  "hypernym": "écoles d'anglais",
  "hyponyms": [
    "cours particuliers d'anglais à Rennes",
    "cours d'anglais en ligne à Rennes",
    "formation anglais à Rennes",
    "cours anglais CPF à Rennes",
    "cours d'anglais pour adultes",
    ... (11 en total)
  ],
  "synonyms": [
    "école d'anglais à Rennes",
    "institut de langues à Rennes",
    "centre linguistique à Rennes",
    "établissement linguistique à Rennes",
    "académie de langues à Rennes"
  ]
}
```

hypernym

Operación 1 (campo léxico) y referente de la operación 3 (anáforas).

hyponyms

Operación 1: al menos 3 deben aparecer en el texto.

synonyms

Operación 2 (sinonimia) y operación 3 (retomas nominales).

Reseñas provenientes de Google Maps vía SerpAPI



Lutèce Langues

4.8/5 ★★★★★ (574) · Academia de idiomas
6 Rue las Cases · Cerrado · Abre a las 9 a.m.



Centre Privé de Langues LSI PARIS

4.8/5 ★★★★★ (202) · Academia de idiomas
23 R. du Renard · Cerrado · Abre a las 9 a.m.



Language Appart'

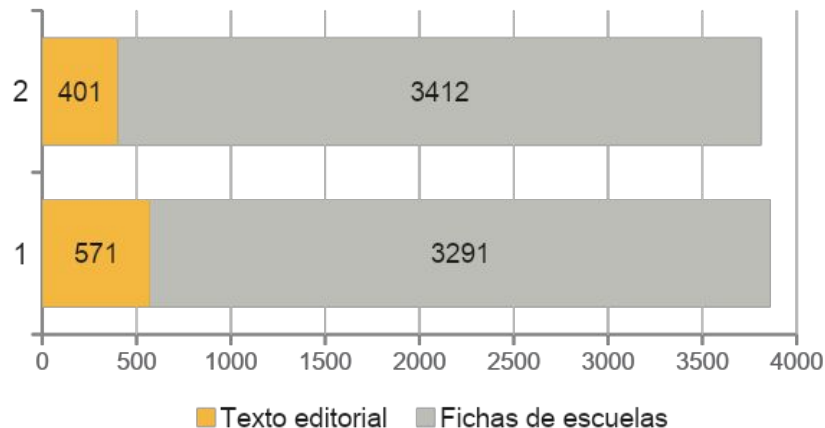
5.0/5 ★★★★★ (333) · Academia de idiomas
10 Rue Gambey · Cerrado · Abre a las 9:30 a.m.

13 ciudades. 130 escuelas. 951 reseñas.

La creación del texto para las páginas de ciudades

```
# H1 (A: reescrito · B: plantilla)
  intro
## H2 por qué aprender inglés en {ciudad}
## H2 criterios para elegir
  ### Qualité des enseignants (solo A)
  ### Diversité des cours (solo A)
  ### Emplacement et accessibilité (solo A)
## H2 selección de escuelas
  ### Escuela 1 ... Escuela 10 ← fichas
  RATING · ADRESSE · SITE
  descripción 120-150 palabras
  POINTS FORTS / FAIBLES
  3 AVIS CLIENTS + enlace Maps
## H2 recomendación final
```

Palabras medias por artículo (11 ciudades)



De 3.860 en promedio, El PES actúa sobre 570 palabras. El 85 % restante del Corpus (A) y el 90 % del Corpus (B) del artículo son fichas de reviews con los mismos datos.

Qué datos usamos y de dónde salen

Herramienta	Bing Webmaster Tools · AI Performance
Qué cuenta	Citas de cada web en respuestas de Microsoft Copilot y partners
Periodo	13 jun – 22 sep 2026 (vista de 3 meses)
Vistas usadas	Total, gráfica diaria, citas por URL y rangos personalizados
Webs	CDB · commentdevenirbilingue.com (sin PES) - AUL · apprendreunelangue.com (con PES)



¿Cómo lo cita Bing Webmaster Tools?

Primer dato: el sitio sin optimización supera 4 a 1



326

citas · sin PES



81

citas · con PES

25 %

AUL tiene la cuarta parte de las citas de CDB

Dónde caen las citas: pocas páginas concentran casi todo

Página	CDB	AUL	AUL / CDB
/ecoles-de-langues/	120	51	43 %
Home	35	0	0 %
París	118	14	12 %
Otras 12 ciudades	48	16	33 %
Total (vista 3 meses)	321	81	25 %

Lectura

- Las páginas con más citaciones son las agrupadoras: suman 155 citas en CDB y 51 en AUL.
- París sola aporta 118 citas a CDB.
- AUL sólo tiene citas en 5 ciudades, las más antiguas; todas sus páginas posteriores al 14 de julio están a cero.



¿Cómo lo cita Perplexity?

El no optimizado suma más citas



13

citas · 4 de 21 respuestas



6

citas · 3 de 21 respuestas

0 de 21

respuestas con los dos sitios a la vez

¿Quién se lleva las citas? 438 citas en 45 dominios

#	Dominio	Qué es	Respuestas (de 21)	Citas	% del total
1	trustpilot.com	Plataforma de reseñas	7	54	12,3 %
2	meilleuresformationsfrancaises.fr	Ranking / blog	10	48	11 %
3	papora.com	Comparador / ranking	9	44	10 %
4	wallstreetenglish.fr	Escuela	6	37	8,4 %
5	britishcouncil.fr	Escuela	6	18	4,1 %
6	commentchoisir.com	Guía de compra	3	17	3,9 %
7	ef.fr	Escuela	6	15	3,4 %
8	cercledeslangues.com	Escuela online	7	14	3,2 %
9	commentdevenirbilingue.com	CDB · estándar	4	13	3 %
10	voscours.fr	Plataforma / blog	8	12	2,7 %
11	englishlive.ef.com	Escuela online	4	12	2,7 %
12	college-council.com	Guía de estudios	2	10	2,3 %
...	...	...	...	...	...
22	apprendreunelangue.com	AUL · PES	3	6	1,4 %



Lo que nos dejan los dos primeros relatos

RELATO 1 · EL REINO DEL CAOS

43,7 %

de las URLs de la U. de la Salud no tienen un tema identificable.

El contenido existe, pero sin clasificación.

+

RELATO 2 · LA PÁGINA QUE NADIE QUISO ESCRIBIR

63 %

de las citas de Bing a AUL fueron a su página agrupadora:

La IA citó la página agrupadora..

ENTONCES

**El problema no es escribir más.
Es agrupar lo que ya existe.**



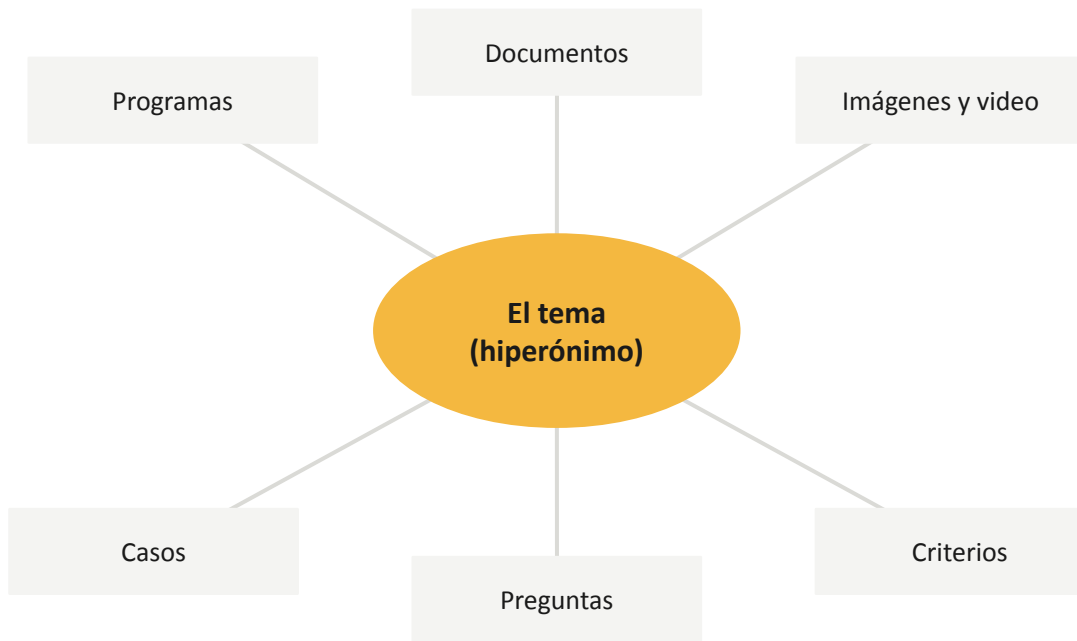
**Necesitamos crear un modelo agrupador
para inspirar**



Relato 3

Domar el caos con una estrategia de contenidos basada en páginas agrupadoras

Una URL con nombre propio que reúne y explica sus piezas



1

Nombra el conjunto

Un título y una definición que dicen de qué trata el grupo.

2

Reúne las piezas

Programas, documentos, imágenes y casos que hoy están dispersos.

3

Explica el criterio

Por qué esas piezas van juntas y cómo compararlas.

Es el PES aplicado a la arquitectura: un hiperónimo y sus hipónimos, explícitos.

SEO y generación de leads para la industria educativa



1

tema

cómo captar estudiantes

10

secciones

una por formato

1

URL

que reúne todo

Una universidad no busca «un caso» ni «un PDF». Busca cómo captar más estudiantes. El tema es el hiperónimo; los formatos, sus piezas.

La misma lógica, aplicada a nuestra propia web.
Casos, guías, artículos y videos, reunidos en una sola página.

Clasificar los formatos según tipologías o temáticas

Carreras de ciencias de la salud

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver las 18 carreras

Medicina	Enfermería	Fisioterapia	Instrumentación quirúrgica
Formación clínica, atención al paciente, diagnóstico y tratamiento de enfermedades.	Atención al paciente, gestión clínica y salud comunitaria.	Rehabilitación del movimiento y a actividades.	10 años de formación, con prácticas.
30 semestres	Preseccional	3 semestres	Preseccional
X Ciencias	Preseccional	X Ciencias	Preseccional

ESPECIALIZACIONES POR ÁREA

Oncología y subespecialidades	Pediatría y neonatología	Medicina interna
Grupos de atención, diagnóstico, tratamiento y seguimiento.	Atención, diagnóstico, tratamiento y seguimiento.	Atención, diagnóstico, tratamiento y seguimiento.
3 años	3 años	3 años

Carreras

Comparador

Compara las carreras de salud

Las mismas variables para todas las carreras, en una sola tabla.

Ver las 12 carreras

Carrera	Duración	Formato de estudio	Modalidad	Áreas de especialización
Medicina	12 semestres	En sala	Preseccional	Neurología, cirugía, medicina interna
Enfermería	6 semestres	En sala	Preseccional	Atención al paciente, gestión clínica
Fisioterapia	6 semestres	En sala	Preseccional	Rehabilitación, deporte, salud comunitaria
Instrumentación quirúrgica	6 semestres	En sala	Preseccional	Quirófano, instrumentación quirúrgica

¿Quieres ver más en la guía de carreras de salud?

Comparador

Del pregrado a la especialización

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver las 12 especializaciones

Pregrado	Especialización	Formación
Enfermería	Enfermería en cuidados intensivos	3 años
Medicina	Medicina interna	3 años
Medicina	Medicina interna	3 años
Medicina	Medicina interna	3 años

Rutas de formación

Videos para conocer las carreras por dentro

¿Quieres y prácticas en YouTube, el día a día de estudiantes en TikTok. Agrupados por carrera.

Ver las 18 carreras

Medicina	Enfermería	Fisioterapia	Instrumentación quirúrgica
¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?	¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?	¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?	¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

YouTube y TikTok

Documentos para decidir con datos

¿Qué futuro tiene un formato que dice lo que contiene, su número de páginas y un formato?

Ver las 8 guías

Guía de carreras de salud	Plan de estudios de Medicina	Guía de administración
Guía de carreras de salud. 100 páginas.	Plan de estudios de Medicina. 100 páginas.	Guía de administración. 100 páginas.

Guías en PDF

Boletines de ciencias de la salud

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver todas las boletines

Boletín de investigación	Boletín de enfermería	Boletín de fisioterapia
Boletín de investigación. 100 páginas.	Boletín de enfermería. 100 páginas.	Boletín de fisioterapia. 100 páginas.

Boletines

Lo que dicen los medios de nuestras carreras

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver todas las guías

Guía de carreras de salud	Plan de estudios de Medicina	Guía de administración
Guía de carreras de salud. 100 páginas.	Plan de estudios de Medicina. 100 páginas.	Guía de administración. 100 páginas.

En los medios

Lo que han logrado nuestros estudiantes

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver todas las guías

Guía de carreras de salud	Plan de estudios de Medicina	Guía de administración
Guía de carreras de salud. 100 páginas.	Plan de estudios de Medicina. 100 páginas.	Guía de administración. 100 páginas.

Premios

Investigados y docentes en LinkedIn

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver todas las guías

Guía de carreras de salud	Plan de estudios de Medicina	Guía de administración
Guía de carreras de salud. 100 páginas.	Plan de estudios de Medicina. 100 páginas.	Guía de administración. 100 páginas.

LinkedIn

Más formas de conocer el área

¿Qué futuro tiene plan de estudios, perfil del egresado, costos y fechas de inscripción?

Ver todas las guías

Guía de carreras de salud	Plan de estudios de Medicina	Guía de administración
Guía de carreras de salud. 100 páginas.	Plan de estudios de Medicina. 100 páginas.	Guía de administración. 100 páginas.

Otros recursos

La matriz tópico × formato: qué existe, qué rehacer, qué crear

Tópico (página agrupadora)	Página	PDF	Infografía	Video	Caso	Preguntas
Especializaciones médico-quirúrgicas	✓	↻	↻	+	+	+
Costos, becas y financiación	↻	↻	+	+	+	+
Investigación en salud pública	✓	✓	↻	+	✓	+
Educación continua en salud	✓	↻	↻	+	+	+
Bienestar y vida en el campus	✓	+	↻	+	+	+



Existe: se enlaza



Existe mal: se reconstruye (nombre, título, carpeta)



No existe: se crea o se trae de fuera (YouTube, LinkedIn)

Cómo se construye la arquitectura, en ocho pasos

1 Inventariar

Rastrear todo: páginas, PDF, imágenes, videos

2 Clasificar

Niveles 1, 2 y 3: tipo, grupo y categoría

3 Definir categorías

Desde las preguntas del aspirante

4 Definir tópicos

Donde ya hay masa de contenido

5 Llenar la matriz

Tópico × formato: existe, rehacer, crear

6 Una casa por pieza

Cada contenido pertenece a un solo tópico

7 Publicar y enlazar

La página agrupadora y sus enlaces de ida y vuelta

8 Medir

Citas en IA y consultas en buscadores

“¿Pensando en estudiar una carrera de ciencias de la salud?”



- Nombre, definición, criterios
- Carreras
- Comparador
- Rutas de formación
- YouTube y TikTok
- Guías en PDF
- Boletines
- En los medios
- Premios de estudiantes
- LinkedIn
- Otros recursos
- Preguntas

1

tema

la pregunta del aspirante

10

secciones

una por formato

1

URL

que reúne todo

El aspirante no busca «un PDF» ni «un video». Busca estudiar salud. El tema es el hiperónimo; los formatos, sus piezas.

**Se agrupa por tema, no por formato.
La web actual lo hace al revés: una carpeta por tipo de archivo.**

Casi todo ya existe: la página agrupadora lo reúne

Sección	Hoy en el rastreo		Acción
Carreras	14 pregrados y 72 especializaciones, cada uno con su URL	✓	Enlazar
Comparador y rutas	Los datos están en las fichas; la vista no existe	↺	Construir desde las fichas
Guías en PDF	42 PDF de costos y admisión, con nombres de archivo opacos	↺	Renombrar y resumir
Boletines	39 URLs, repartidas en egresados, internacionalización y proyección social	↺	Reunir por área
Premios	29 URLs de reconocimientos, entre noticias y fichas de programa	↺	Reunir
YouTube y TikTok	23 URLs de video en el sitio; los canales, fuera	+	Incrustar por carrera
En los medios	0 URLs de prensa en el rastreo	+	Crear y enlazar a la fuente
LinkedIn	Fuera del sitio	+	Enlazar publicaciones
Otros recursos	No existen como sección	+	Crear

6 de 10 secciones salen de lo que ya está en el sitio. Lo nuevo es el orden.

Boletines de la Universidad de la Salud

Universidad de la Salud

Boletines de la Universidad de la Salud

3.836 boletines

5 años

8 años de historia

Filtros: año, público, facultad

Última edición, resumida en HTML

Cinco públicos, cinco series

Histórico por año

Suscripción

- Nombre, definición, tamaño
- Filtros: año, público, facultad
- Última edición, resumida en HTML
- Cinco públicos, cinco series
- Histórico por año
- Suscripción

3.836

boletines

en una sola URL

5

públicos

una serie para cada uno

8

años

de histórico

Un formato que se publica cada mes necesita dos cosas: un lugar donde encontrar lo último y un archivo donde volver a lo anterior.

El público organiza. El año archiva.



**La respuesta depende menos de lo que van
a publicar mañana y más de cómo ordenen
lo que ya publicaron.**



¡Gracias!

David Martínez Mata

Máster Hispanofrancés en Lengua Francesa Aplicada, Universidad Complutense de Madrid.
Master Sciences du Langage - Français Langue Étrangère, Sorbonne Université, Paris.