

Recurso Educativo Abierto: Trabajando con datos de los Sistemas de Portales de Obligaciones de Transparencia del gobierno de México en educación

Víctor Manuel Martínez Martínez
vicm3@upn.mx



Porque este tema

- Licenciatura en Administración Educativa
- Originalmente como idea para [Edusol 2017](#), capítulo de libro
- Nos dejó con dudas y pensando más en el tema
 - Qué se puede mejorar
 - Qué encontramos
 - Qué nos faltó



Año sabático 2019

- Motivación principal
- Probar y equivocarse
 - Big Data, Web Scraping, ML
 - Abordar de forma equivocada el tema (en parte)
- Errar es de humanos, herrar de herreros
 - Pero encontramos el camino
- Construir un material didáctico, que sea un REA



Que queremos

- Que los estudiantes y nosotros podamos obtener datos sobre la propia Universidad
 - Matricula, egresados, ingreso, reingreso, becas, etc.
- Tarea terriblemente tediosa y mecánica
- ¿Lo podrá hacer una maquina?



Descubrimos

- La institución y su portal(es) de Obligaciones de Transparencia
- El Portal de Obligaciones de Transparencia y el ahora Sistema de Portales de Obligaciones de Transparencia (SIPOT)
- Quien diseña el sistema en el Instituto Nacional de Transparencia (INAI) no es el mismo que termina llenando los reportes en el organismo obligado
- Peor aún la implementación del portal institucional normalmente la hace un tercero que no tiene relación con los dos de arriba
 - Que la dificultad de consulta no es debido a malicia, es prisa por cumplir con las obligaciones
 - Que mucha información esta disponible para cumplir las obligaciones, no porque sea útil



Horrores e técnicos

- De parte de UPN hay:
 - <http://transparencia.ajusco.upn.mx> PHPNuke
 - <http://sipot.upnvirtual.edu.mx> Joomla
- De parte de INAI
 - <http://portaltransparencia.gob.mx>
 - <https://consultapublicamx.inai.org.mx>
- !Todos son diferentes! Y en ocasiones un sistema te lleva a otro o te avisa que la información la puedes encontrar en el otro sistema, ejemplo: nuevo ingreso



Horrores técnicos...

- UPN usa Joomla para sus portales y phocadownload, genera URL como este:
 - http://sipot.upnvirtual.edu.mx/index.php?option=com_phocadownload&view=category&download=1940:se14&id=381:2018&Itemid=310
- Todos los documentos son PDF
- Por eso Google y demás no nos ayuda, las páginas son generadas “dinamicamente” y no tienen un destino real, fuerzan la descarga, entonces no se encuentran en los índices de búsqueda
 - Solución
 - PDF → xpdf/pdftotext → Texto plano
 - PDF → xpdf/pdftoimages → imagen → gocr → Texto plano
 - En el escritorio Recall



No se puede con ML/Big Data

- <<“**Big data**” often means “**data produced by someone else**” and there’s lots of it>> Tom Smith (2019)
- Sociólogo que puede hacer bastantes cosas con la computadora
 - “Cliff, he’s not much of an astronomer, but what a computer hacker [...] Cliff’s not much of a programmer, but what an astronomer!”, The Cuckoo’s Egg
- Trate con xpdf → Texto → Wordcloud, problema, “no se habla español”
 - Seminario de redes IIEc (gran hack)
- Sin usar herramientas de la nube, porque la nube, **es la computadora de alguien más**
- Mi trabajo es enseñar a futuros Administradores Educativos a obtener los datos y la información, el nivel de complejidad resulto muy alto
 - No es que no se pueda, yo no puedo en el estado actual y no encontré las herramientas para hacerlo posible para un estudiante de LAE UPN
 - Hackear la cosa de ML a Aprendizaje humano



Recurso Educativo Abierto (REA)

- Por que los profesores publicamos poco o nada
- Profesores que dan excelentes clases pero no se sientan a escribir (mea culpa)
- De Socrates conocemos por uno de sus alumnos, Platón que escribió grandes cosas de su maestro
 - No publico un texto
 - Pocos estudiantes
 - Exámenes orales
 - Hoy no seria candidato a SNI/PRODEP
- El sabático se paga con dinero público, que los resultados sean públicos
 - Cola de espera para publicación UPN de varios años
 - Archivo de estímulos y becas
 - Rincón de la biblioteca
 - Otra forma de que quien necesita buscar en transparencia lo pueda lograr



La solución

- Encontrar el sistema de gobierno de UPN u otra IE
- Quien reporta, en este caso organismo desconcentrado, la Comisión Interna de Administración (CIA)
- Gobierno, ciclos trimestrales
 - Resultado, cuatro informes al año, de 2004 a 2019, es decir 60 informes...
- El objetivo era fortalecer contenidos de dos materias, se nos ocurrió una materia optativa



REA

- Aprendizaje basado en problemas
- Queremos que el estudiante razone que busca, no darle una receta
- Ejercicios que den pistas y no la respuesta
- ExeLearning para autoría
- Pendiente
 - Rubricas, métricas, test pre / post (aceptamos sugerencias)



Lo faltante

- Datos abiertos
 - Gran idea, pésima implementación, en proceso de abandonado
- Si hay lineamientos internacionales [OEA](#)
- Cómo se asegura
 - Con sanciones económicas
 - Revisión aleatoria INAI
 - Denuncias personales y anónimas
- ML/AI/etc. Contacté con un colega experto en el tema, que ha dado luz para un enfoque diferente, falta probar...



Gracias por su atención

- ¿Preguntas, comentarios?

Víctor Manuel Martínez Martínez

vicm3@upn.mx

