

# J'ai demandé mes données personnelles grâce au RGPD

Antoine Augusti

Licence [CC BY-SA](#)

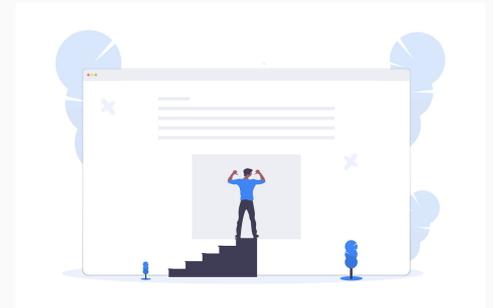
# Qui suis-je

Salut ! Je suis Antoine Augusti.

**Formation** : ingénieur en informatique INSA Rouen  
Normandie, master de recherche en statistiques

**Professionnel** : software engineer @ Etalab ( Direction interministérielle du numérique / services du Premier ministre)

**Contact** : [@AntoineAugusti](https://twitter.com/AntoineAugusti) / [www.antoine-augusti.fr](http://www.antoine-augusti.fr)



# Droit à la portabilité

Spécifié dans l'article 20 du RGPD

Les personnes concernées ont le droit de recevoir les données à caractère personnel les concernant qu'elles ont fournies à un responsable du traitement, dans un format structuré, couramment utilisé et lisible par machine et ont le droit de transmettre ces données à un autre responsable du traitement.

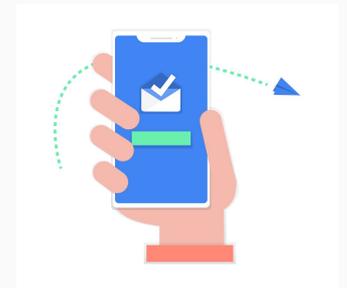
# Exercer son droit à la portabilité

Utiliser un outil en ligne (dans un monde idéal) ou suivre la procédure décrite dans les CGU ou la politique de confidentialité.

Bonjour,

Je souhaite exercer mon droit à la portabilité de mes données en vertu du RGPD. Merci de m'envoyer ces données par e-mail dans un format ouvert. Si vous possédez une documentation associée, merci de me la communiquer également.

En l'absence de réponse dans un délai de 30 jours, je contacterai la CNIL.



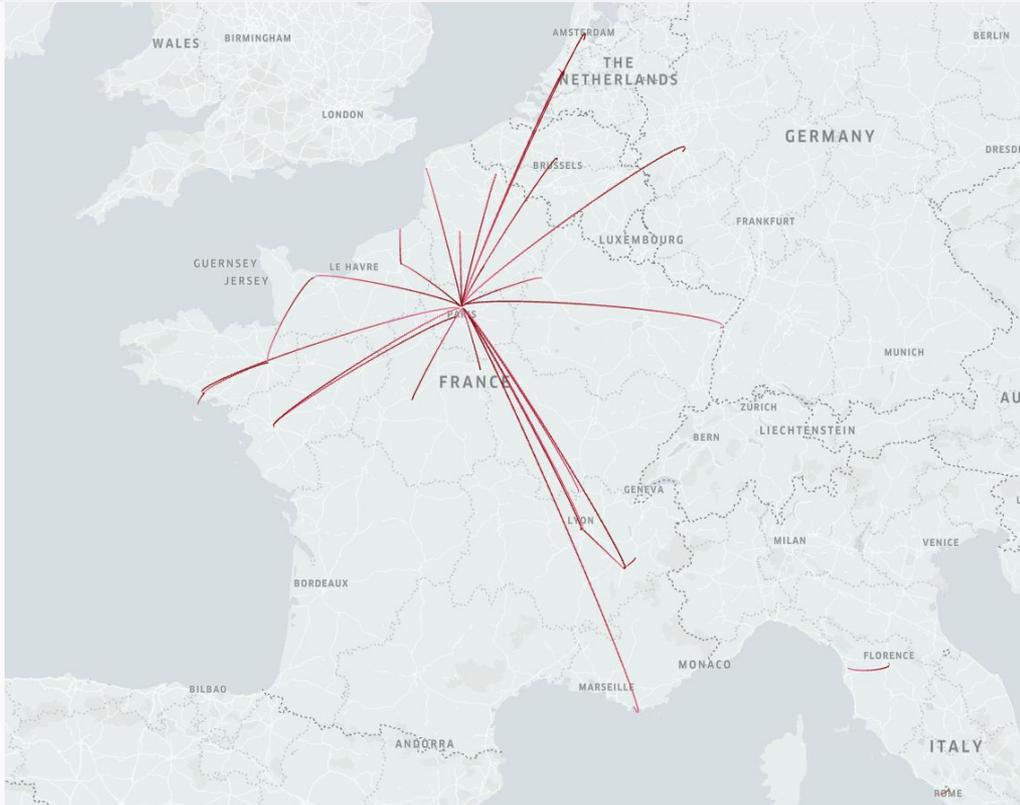
# Visualiser mes trajets en train grâce à Trainline

Trainline (anciennement Capitaine Train) vend des billets de train en Europe. J'utilise ceci depuis plus de 6 ans.

Suite à ma demande, j'ai reçu un **unique fichier JSON de plus de 110 000 lignes** (données du compte, passagers, cartes de crédits, voyages, recherches), sans documentation.

1. Nettoyage des données en Python ([AntoineAugusti/trainline-gdpr](#))
2. Visualisation des données en utilisant une carte interactive ([kepler.gl](#) de Uber)

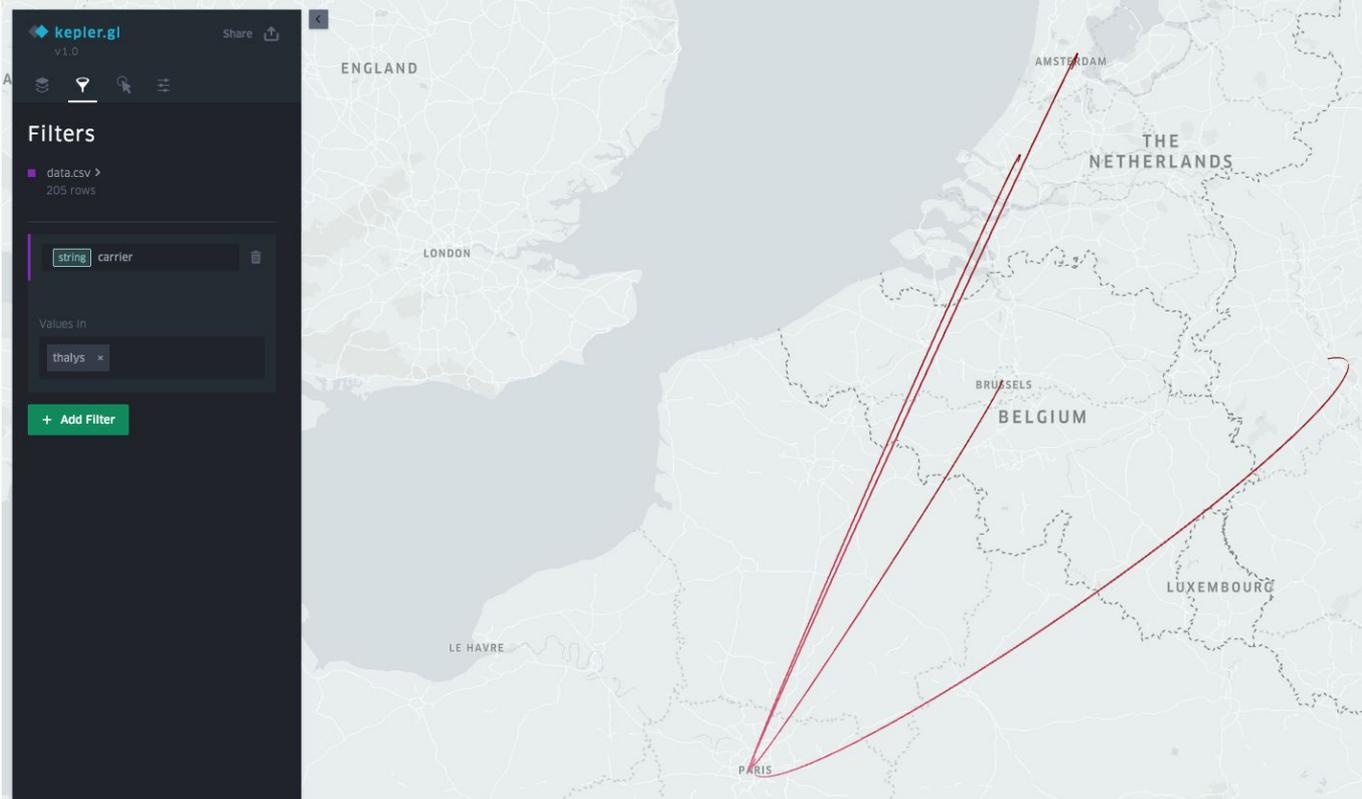
# Tous mes trajets en Europe (presque)



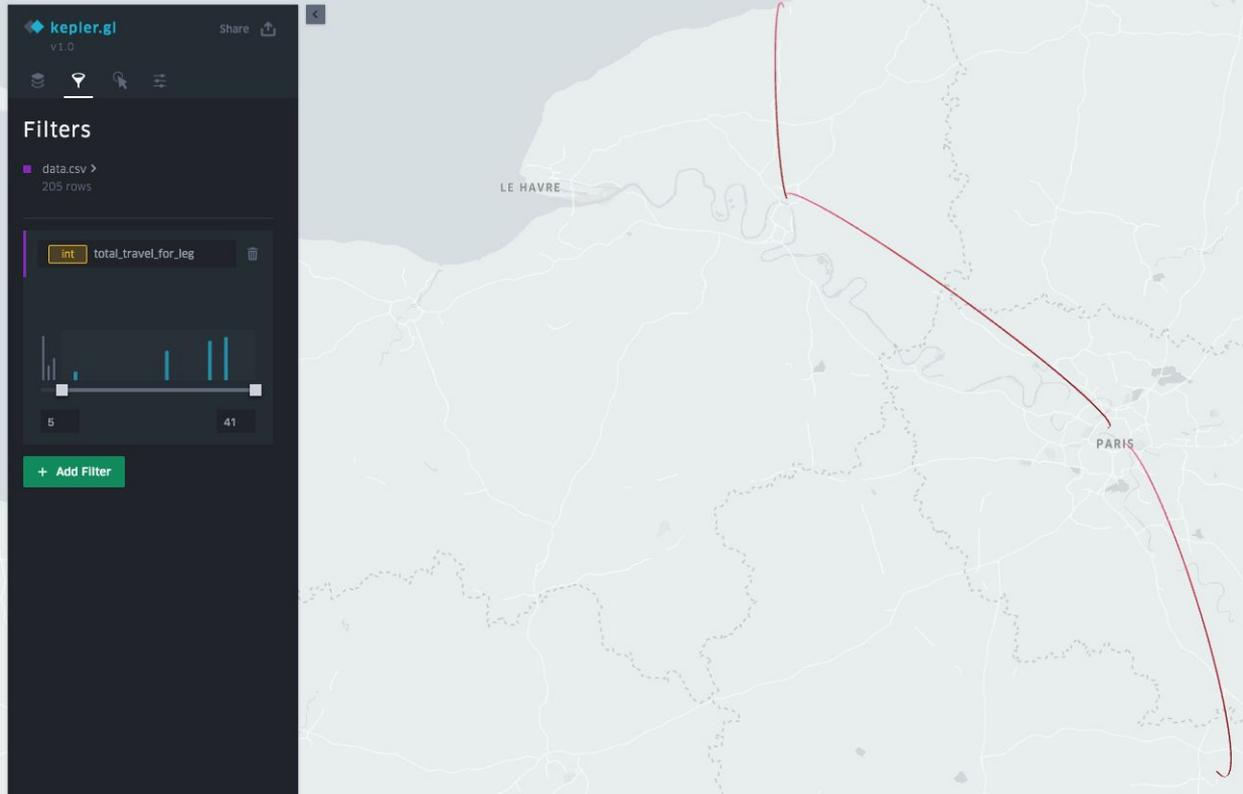
## Données manquantes :

- Billets que l'on ne peut pas acheter sur Trainline (achat avec des machines)
- Billets pour des trajets professionnels
- Billets Eurostar (moins chers en £)
- Billets à l'étranger où je préférerais parler à un agent pour avoir des informations

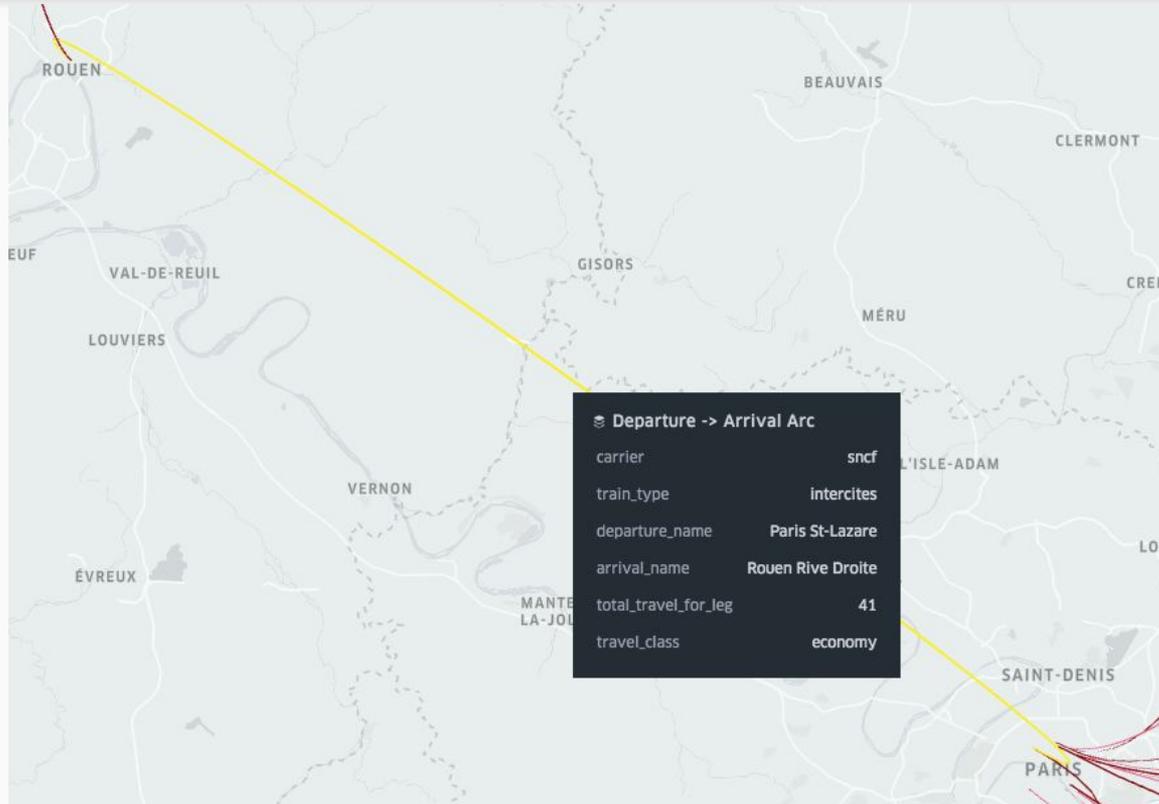
# Trajets en Thalys



# Trajets que j'ai fait plus de 5 fois



# Trajet le plus commun



# Portabilité des données Monoprix

J'ai récupéré mes données d'achat de Monoprix par le biais de ma carte fidélité.

Nos habitudes :

- Nous mangeons végétarien et ne buvons pas d'alcool à la maison ;
- Nous achetons des aliments végétariens et fruits et légumes ailleurs ; pain à la boulangerie
- Notre Monoprix est situé au centre de notre trajet domicile - gare RER, qui prend moins de 5mn à pieds ;
- Les données vont de début 2017-01 à fin 2018-09.



# Données Monoprix

A	B	C	D	E
Day of week	Number of visits	Amount spent	Mean of amount spent	Mean of local time
1	48	561.03	8.62	18
2	31	450.2	9.44	18
3	35	258.51	4.99	18
4	16	147.48	6.78	18
5	23	277.4	8.18	18
6	38	501.15	11.275	12
7	13	211.41	19.2	11
<b>Grand Total</b>	<b>204</b>	<b>2407.18</b>	<b>8.02</b>	<b>18</b>

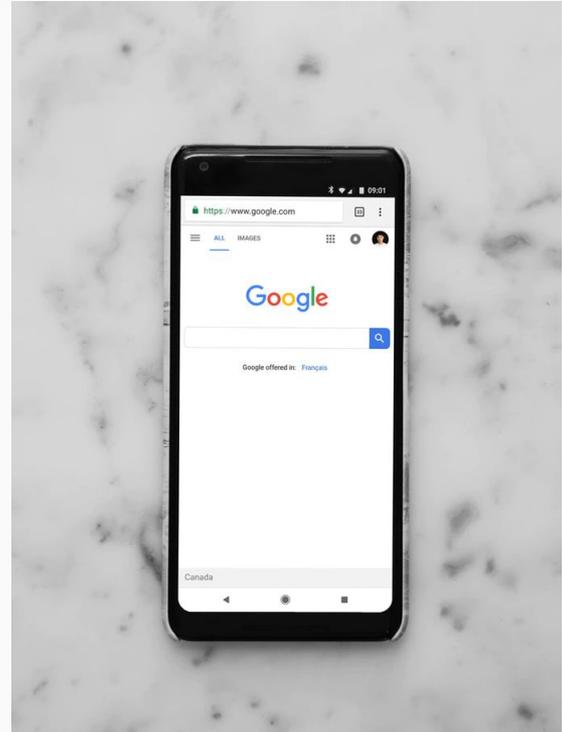
Year and month	Number of visits	Mean of amount spent	Total spent	Median hour of visit	Mean of week day
2017-01	3	6.16	53.78	20	2
2017-02	10	8.67	115.01	18	2.5
2017-03	11	7.48	139.09	19	2
2017-04	11	14.54	209.59	18	3
2017-05	11	8.25	152.97	18	3
2017-06	11	7.28	119.58	18	3
2017-07	12	10.195	147.69	15	3.5
2017-08	9	7.01	97.96	19	2
2017-09	4	6.505	30.79	19	2.5
2017-10	10	11.115	111.3	14.5	4.5
2017-11	11	9.19	113.17	13	6
2017-12	9	7.33	104.61	18	4
2018-01	9	9.93	114.35	18	3
2018-02	11	9.33	133.83	17	4
2018-03	10	7.76	138.2	17.5	4
2018-04	14	6.28	122.84	18	3
2018-05	9	10.9	104.11	18	3
2018-06	12	6.32	139.86	18	3.5
2018-07	9	6.96	86.13	18	3
2018-08	9	7.04	72.19	18	3
2018-09	9	9.48	100.13	18	3
<b>Grand Total</b>	<b>204</b>	<b>8.02</b>	<b>2407.18</b>	<b>18</b>	<b>3</b>

# Données Google Search

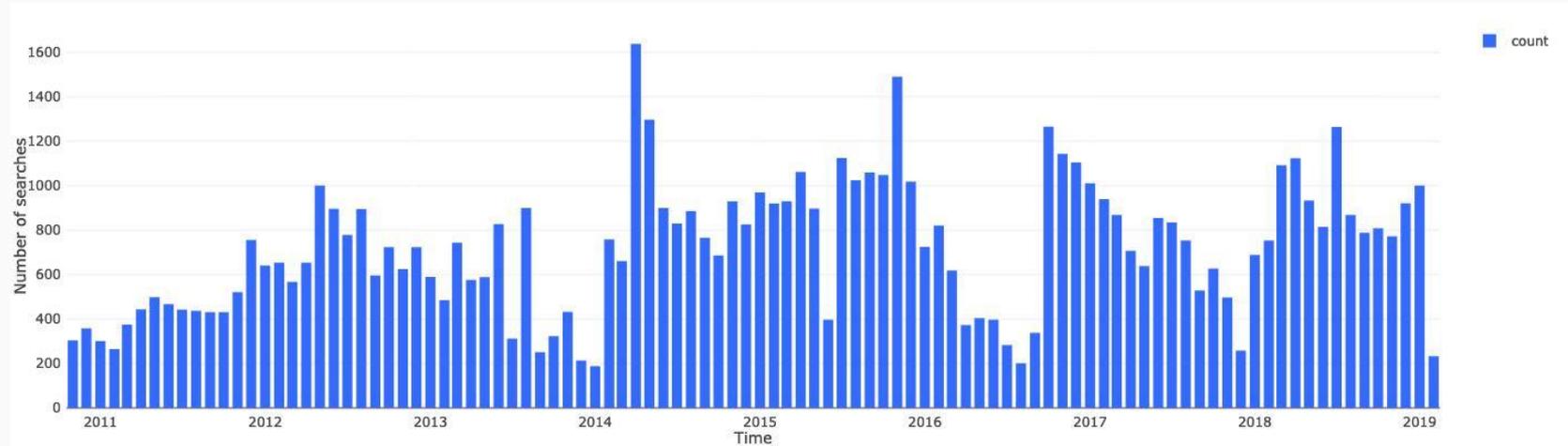
J'ai exporté mes recherches sur Google où j'étais connecté avec mon compte. 71 600 recherches depuis 2011.

Moyenne de 25 recherches / jour, tous les jours, pendant 8 ans.

1. Utilisation de [takeout.google.com](https://takeout.google.com)
2. Nettoyage en Python des données [AntoineAugusti/google-search-gdpr](https://antoinerobert.com/antoinerobert/google-search-gdpr)



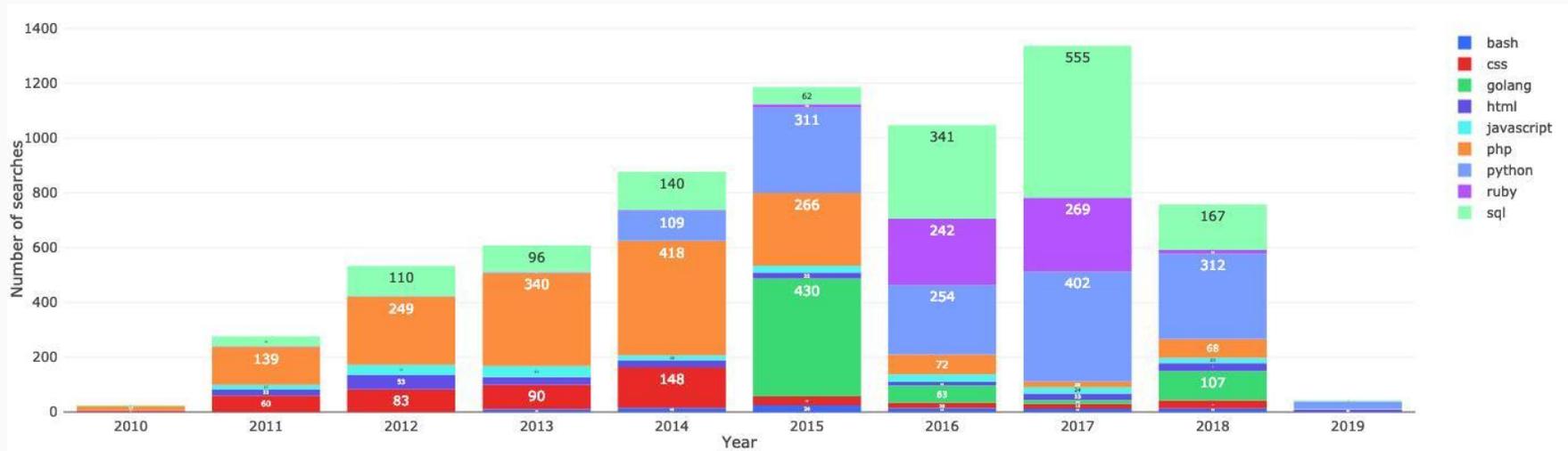
# Recherches par mois



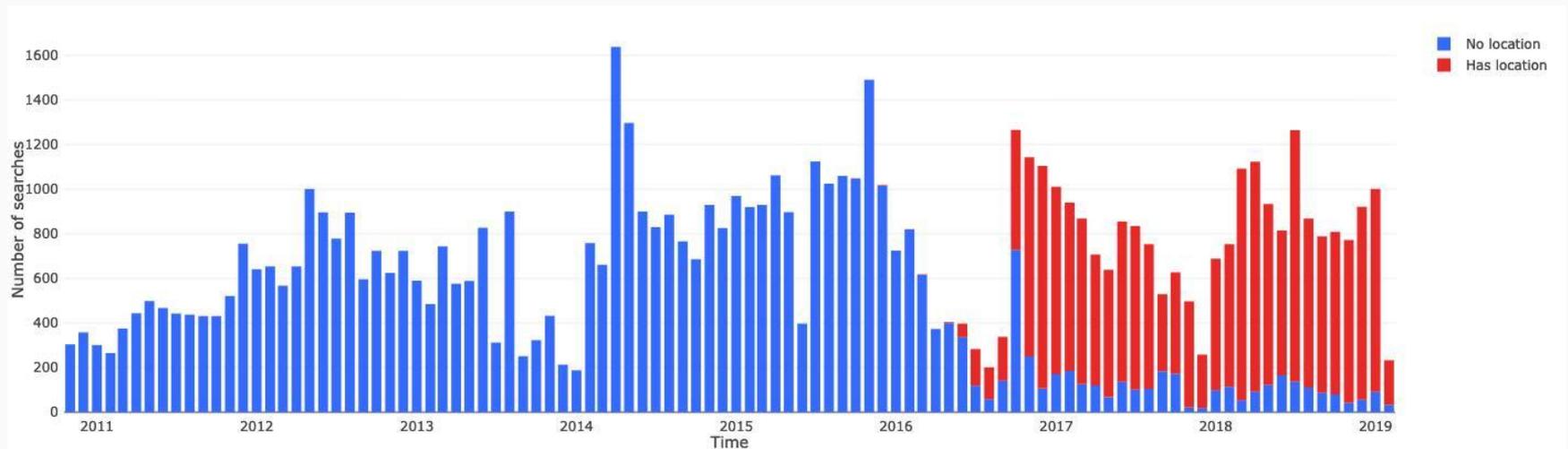
# Recherches par heure

weekday	hour	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	Totals
<b>0 - Sunday</b>		21	11	19	4	5	19	174	391	624	796	924	716	659	732	651	680	827	1,016	754	633	555	353	105	29	<b>10,690</b>
<b>1 - Monday</b>		2	1	6	4	19	68	135	411	639	727	571	570	665	740	714	679	668	722	642	640	629	376	111	53	<b>9,847</b>
<b>2 - Tuesday</b>		4	11	9	3	19	44	215	461	812	896	684	527	636	738	701	661	621	852	645	646	588	405	98	18	<b>10,294</b>
<b>3 - Wednesday</b>		6	5	13		19	66	177	462	792	781	594	527	568	645	706	688	668	751	682	667	661	377	151	21	<b>10,027</b>
<b>4 - Thursday</b>		14	9	6	1	14	82	239	326	662	734	588	555	633	792	801	682	854	838	602	578	512	392	209	51	<b>10,172</b>
<b>5 - Friday</b>		11	4	4		22	75	218	407	713	788	656	535	691	669	582	662	687	815	608	624	684	498	254	78	<b>10,285</b>
<b>6 - Saturday</b>		18	16	16	5	13	24	249	400	576	668	750	650	686	568	547	673	809	915	563	573	547	642	284	66	<b>10,254</b>
<b>Totals</b>		<b>76</b>	<b>57</b>	<b>73</b>	<b>17</b>	<b>111</b>	<b>378</b>	<b>1,407</b>	<b>2,858</b>	<b>4,818</b>	<b>5,390</b>	<b>4,767</b>	<b>4,079</b>	<b>4,538</b>	<b>4,884</b>	<b>4,702</b>	<b>4,725</b>	<b>5,134</b>	<b>6,960</b>	<b>4,496</b>	<b>4,361</b>	<b>4,176</b>	<b>3,043</b>	<b>1,212</b>	<b>316</b>	<b>71,578</b>

# Recherches par langages de programmation



# Recherches avec localisation enregistrée



# Problèmes actuels avec la portabilité

- Très peu d'entreprises ont des procédures pour gérer les demandes de portabilité. Peu de personnes demandent leurs données ;
- Comment pouvez-vous être sûr que l'on vous donne bien toutes vos données ;
- Manque de documentation, ambiguïté des termes utilisés ;
- Besoin d'outils pour nettoyer et analyser ce qu'on obtient : bibliothèques, interfaces web, articles de blog. Actuellement abordable seulement avec des connaissances techniques.
- Pour exercer pleinement son droit de portabilité, il faut des structures de données communes entre des services similaires. Data Transfer Project.

La fin

**Merci !**

Questions ?