Reuters Institute Fellowship Paper

University of Oxford

How media companies use data to sign up digital subscribers (and keep them)

by Michael Leitner

Michaelmas 2018

Sponsor: APA (Austrian Press Agency)

Table of contents

Acknowledgements

I would like to thank the Reuters Institute for the Study of Journalism and the Austrian Press Agency for giving me the chance to spend three months as a journalist fellow in Oxford, where I was able to gain new knowledge and perspectives, meet fascinating people from all over the world and work on my research. I owe the most thanks to my supervisor Anita Zielina, who supported and impressed me with her extensive knowledge on paywalls and digital media, as well as Meera Selva and Rasmus Kleis Nielsen, who provided me with valuable feedback on my research and opportunities to meet executives from the media industry.

I would also like to express special gratitude to my friends and former journalist fellows Christoph Schlemmer and Alexander Fanta, who inspired and supported me in applying and preparing for this fellowship program. Special thanks are due to my hosts Janie and Charles Hampton as well as my housemate Helen Jeans, who made me feel at home in Oxford. And last, but not least, I want to thank everyone – family, friends and colleagues – who supported me in the months before, during and after my time in Oxford. Becoming a journalist was already a dream come true for me, getting the chance to research at the University at Oxford was another one.

1. Introduction

"You have read your ten free articles this week. Please subscribe to continue reading." Almost every internet user probably hit the limit on a so-called metered paywall at least once. The idea is simple: You can consume a specific number of articles for free, after that you have to buy a digital subscription for further access. When the New York Times announced this model for their web content in 2010, they were met with scepticism by readers and industry observers alike (Pérez-Peña, 2010; Rozvar, 2011). US blogger and author Cory Doctorow even called it crazy, hard to understand and potentially frustrating (Doctorow, 2011). Nevertheless, many legacy media organizations followed suit and adopted a paywall, most of them based on a metered model (Williams, 2016).

While the metered model turned out to be a success for the New York Times, other organizations struggled with it. Some considered this to be proof that digital users are still not willing to pay for news and abandoned paywalls altogether, but the majority accepted there's no "one size fits all"-solution and kept adapting their strategies. Today, most media organizations with a paywall use the so-called freemium model, which puts only specific articles behind the paywall (The Newsplexer Project et al., 2018; Cornia, Nielsen et al., 2017). In some way this sounds like a return to the beginning of paywalls, when media organizations began to charge for content they considered to be 'premium'. But this time, it is successful.

Smart and user-focused

One of the reasons for this success: Media organizations started listening to their users' needs. Modern paywalls are data-driven, so only content that efficiently convert readers to subscribers will be put behind the paywall. Many well-known media organizations such as the Telegraph, Die Welt and the Süddeutsche Zeitung switched successfully from a 'dumb' rule-based to a 'smart' data-driven model, (Sweney, 2016; Meedia, 2016).

This 'smart' approach usually has data at its core, but the way this data gets collected, processed and applied can vary strongly from organization to organization. While some create algorithms that personalize what content stays behind the paywall based on the users' behaviour, others prefer a 'human' approach and create easy-to-understand metrics, that help them quickly steer their editorial strategy based on their subscribers' interests (Piechota, 2018). Many organizations also use a mix of both, for example by personalizing offers and 'Call to Action'-messages with algorithms, and creating metrics not just for the newsroom, but also other departments, such as product development, marketing and sales.

But efficiently working with data requires skilled employees, modern technology and flexible structures – attributes mainly found in tech companies and start-ups. It is no small task for legacy media organizations, who tend to be characterized by bureaucracy, risk aversion and gut-based decisions, to acquire these traits. But it's also a challenge for digital-born media organizations, who often have to work with a limited budget and face other existential challenges. Nevertheless, it is possible, as it has been proven by several successful examples.

1.1. Research questions and methods

This paper aims to look behind the scenes of those media organizations who already use data to find out what's necessary to efficiently acquire, engage and furthermore keep subscribers. It will look at

two specific types - the automated and manual 'smart' paywalls, though this paper will focus on the former. Automated paywalls are ones where an algorithm individually decides what and when content gets put behind the paywall. Manual 'smart' paywalls describe those systems where the content for subscribers still gets chosen manually, but the decision on that is based on data. I aim to compare the two to draw out advantages and disadvantages in the two approaches.

This paper aims not to answer what specific types of content get chosen for paywalls, but rather how media organizations created a process to efficiently pick the right content to acquire and retain as many subscribers as possible and what data they based their decisions on.

Considering that some legacy media organizations, such as newspapers, have already been in the business of selling subscriptions and collected data on their subscribers, this paper also aims to answer the question if those organizations were able to benefit from their past experiences and if it influenced their current digital strategies.

Thus, the main research questions for this paper are as follows:

- *How do media organizations use data to improve the efficiency of their subscription strategy?*
- *What are the advantages and disadvantages of using automated processes for your paywall?*
- *Who is able to use dynamic paywalls at their organization and what are the requirements?*

Only a handful of media organizations already use automated paywalls, so the choice is mostly limited to big and well-known brands, such as NZZ, the Wall Street Journal and the Norwegian media group Schibsted.

There are a broader range of media organizations who use data to manually steer their paywall strategies, but in order to ensure comparability between the two groups, I've picked mostly legacy media organizations, such as Der Spiegel and Die Welt, but also the relatively young Slovakian newspaper Dennik N, which was co-founded by a former executive of Piano Media.

Academic papers which compared the business and editorial strategies of various news organizations with paywalls already exist (Nielsen, Nicholls & Shabbir, 2016; Cornia, Nielsen et al., 2017). This paper aims for a different perspective by focusing on how media organizations make use of metrics and data analysis in order to adapt their strategies to the user. To offer a more vivid image of the process, this paper will also show what kind of specific new products or process improvements they came up with in order to acquire and retain subscribers.

This paper looks at the various parts of the subscription business and therefore starts with everything related to acquisition before moving on to engagement and retention. The acquisition chapter gives a rough overview of the current state of the industry, with a major focus on dynamic paywalls. Because engagement and retention are so closely linked, they share the following chapter, which also has a major focus on dynamic and automated approaches, their functionality and best practice examples. The final chapter explains how the organizations interviewed for this paper are structured, how they had to adapt in order to work efficiently with data and what kind of metrics

and KPIs they use. The final chapter is intended to give a broad summary of the findings and answer the research questions proposed in this chapter.

1.2. Who should read this?

This paper is aimed at anyone at a media company that is planning to offer subscription-only content or already does, especially those involved in product management and audience development.

The number of media organizations who will adopt subscription products is only going to grow. Stricter rules on data collection and privacy, the so-called GDPR (General Data Protection Regulation), are probably going to accelerate the adoption of subscription products by a large factor, though they'll create additional challenges for data collection. This new law requires explicit consent from the user if a website wants to collect data on them. Several news organizations are going to take this challenge as a chance and will offer ad-free subscriptions. Google and Facebook already announced several products that help push subscriptions for publishers (Brown, 2018; Moses, 2018a). Young people, who grew up with subscriptions for entertainment content, such as Netflix or Spotify, seem to be more open to the idea of paying for news and could drive future growth (Fletcher, 2017).

2. Key Insights

Dynamic paywalls are no silver bullet solution: Dynamic paywalls are a new hybrid model (metered and data) that tries to personalise the user experience of the paywall. In some way they are a logical evolution of the current paywall models and their shortcomings. Whereas the metered and freemium model have one set of rules for everyone, the dynamic paywall tries to adapt to the individual user, for example by dynamically adjusting the meter limit or personalising marketing materials. Although the media organisations who already have solutions like this in place are satisfied with its results, they're still no "one size fits all"-solution and come with several trade-offs and a long list of requirements, for example big investments into data science and technology.

If you want to keep your subscribers, watch their behaviour: Behavioural data seems to have the strongest correlation with user satisfaction and therefore retention. Most organisations follow the RFV model (Recency - Frequency - Volume) created by the Financial Times in order to set KPIs for their retention strategies. Based on that they try to increase engagement and retention by forming habits, for example by improving the onboarding process and creating content that gets subscribers to come back regularly. Demographics, content types and other metadata, like operating systems, can be a predictor for subscription likelihood and retention, but aren't as relevant as behaviour.

Predicting churn is easy, preventing it is hard: By tracking the behaviour of subscribers, it's getting easier to actually identify those users who are about to cancel their subscription. But what are you supposed to do with that kind of information? Many organisations interviewed for this paper said even though the data models are quite accurate, they haven't found economic solutions that prevents those users from leaving. Reaching out to users who rarely engaged with their subscription sometimes even had a negative effect, because they were reminded of their ongoing subscription and cancelled it. Instead they're trying to tackle the churn problem by focusing on engagement and forming habits early on, which in long-term increases the retention rate. But the churn prediction models are still in use, especially at customer service and telemarketing departments, who'll try to convince users with a high churn risk to renew or keep up their subscription. This human approach has proven to be more successful than automated measures by mail or email.

Privacy regulation adds a big question mark: New privacy regulation, such as the GDPR (General Data Protection Regulation) and ePrivacy regulation (which is supposed to come in effect in 2019), could hamper the effectiveness and industry adoption of dynamic paywalls. Several publishers voiced their concerns in interviews conducted for this paper, especially about the ePrivacy regulation, which in its current form would prohibit cookie-based tracking without explicit affirmative consent by the user. A possible solution could be to push users to log-in on the respective site, because they usually accept the privacy terms when signing up for an account. But this would still limit the effectiveness of the data model, because only a small share of total users could be tracked.

New players focus on acquisition, first movers bet on retention: Retention still seems to be an afterthought for many publishers who only recently started to offer paid content. They rather focus on acquisition in order to build a reasonable sized audience and to gain an advantage over competitors. This strategy focus is often not limited to a single player, but the whole market. 'Less developed' markets like Germany, where paywalls haven't been as prevalent as in other countries,

tend to have tough competition and a heavy focus on subscriber numbers, whereas in 'mature' markets like Norway or Sweden, the competition softened and organisations tend to focus on revenues and retaining their existing subscribers.

Don't try to hit everything with the machine learning hammer: Although it's tempting to automate as much as possible in the current data-driven environment, it doesn't always make sense to do so. Several publishers interviewed for this paper pointed out that they abandoned data projects when it turned out that either the possible savings didn't justify the costs or when they couldn't find a lever to influence what they were measuring. Schibsted also created a culture where employees still hold the upper hand over algorithms, for example by setting variables like news value and longevity themselves and regularly discussing the effectiveness of the data models with different departments, to address a wariness about "autopilots" and "black boxes". Current studies seem to indicate that the media industry as a whole is cautious about automation. According to a PwC study, only seven percent of interviewed media executives are planning to make "substantial investments" into AI technology in 2019, much less than the general average of 20 percent (Spangler, 2018).

Know who, what and why you're tracking: In order to successfully automate processes, you need to be able to answer several key questions, such as who is your targeted audience as well as what kind of data do you want to collect and why. The latter is especially important when it comes to creating data models, not just for legal reasons, but also to make the impact of your solutions measurable and reduce cost. Media organisations interviewed for this paper used between 10 and 100 different signals for their dynamic paywalls, but tested thousands before actually deploying them. That's why organisations who had data and metrics deeply embedded into their culture were quicker to embrace dynamic paywalls.

The newsroom needs to understand your goals and metrics: In order to make a paywall work, whatever the model may be, you need to have clearly defined goals that everyone in the organisation, especially the newsroom, can understand and work towards to. Most media organisations interviewed for this paper refrained from using aggregated scores and rather focused on simple and easy to understand KPIs, which were aligned with the business goals. The Wall Street Journal for example identified a rather simple metric - active days within the last 28 days - as one of their most crucial KPIs, because it was so strongly connected to churn likelihood and therefore also improved engagement and retention. Organisations also need to provide ways for employees to provide feedback and ask questions about metrics in order to improve "metrics literacy".

Create flexible structures, so everybody can participate: Especially legacy media organisations struggle with complicated hierarchies and loads of red tape, that often slow down change processes. That's why most of the organisations interviewed in this paper created interdisciplinary departments, consisting of data scientists, project managers and other highly skilled employees, who can quickly form small project teams with other employees and stakeholders within the organisation. This can be highly motivating for independent-minded employees who get the chance to be part of a usually non-transparent process and bring their own ideas to the table.

3. Acquisition

When it comes to subscription products, the strategy is usually split into three parts: acquisition, engagement and retention.

Acquisition describes measures required to get a user to become a subscriber, usually structured in a so-called sales funnel. The further the user gets down the funnel, the more likely it is he or she becomes a subscriber.

Engagement describes methods and processes to ensure the subscriber uses the service as much as possible. Because high usage of a service is usually linked to high customer satisfaction, organisations with a high share of subscription revenue tend to invest heavily in engagement measures.

This should further drive retention, which describes measures and processes to ensure a user keeps a subscription active. Measures to stop someone from cancelling their subscription can but don't necessarily involve improving engagement, because reasons for cancellations are more varied than just dissatisfaction with the service, for example pricing or competitors. Retention is a crucial part of subscription businesses, because it determines how much of your previous efforts end up as beneficial to the bottom line.

The whole process is circular: You have to get subscribers in order to see how well your service performs and improve its usage. When users still want to leave, you have to find out why and put countermeasures into place, either to stop them from leaving or others following suit.

3.1. Current state of the industry

Despite this clear definition of the process, many media organisations interviewed for this paper pointed out, that they mainly put their focus on the acquisition process for the first few years. „We didn't care about anything but acquisition for the first three years. And now we are kind of starting slowly to do something in [engagement] ", says Tomáš Bella, one of the co-founders of the Slovakian daily newspaper Denník N. Der Spiegel, who started Spiegel Plus in May 2018, voiced a similar stance: „In the beginning we want as much growth as possible. Fine calibration will happen, when we've reached a ceiling ", says Elisabeth Kanzi, product manager at Spiegel Online. Even Schibsted Media Group, who offered paid content at multiple brands for years, pointed out that they've only recently begun shifting from acquisition to engagement and loyalty metrics.

Even though only six different media organisations were surveyed for this paper, a bigger sample indicates that the whole industry is still in the process of putting more effort into engagement and retention. The INMA (International News Media Association) surveyed 60 companies who were attending their Consumer Engagement Summit in 2018 if they spent more on acquisition or engagement. 59 percent favoured acquisition, but companies who invested more in engagement reported higher satisfaction with the results of their strategy (The Newsplexer Project et al., 2018, p.10).

The focus on one part of the process can be beneficial in the short term, for example to quickly grow a substantial subscriber base, but to ensure lasting success, every aspect needs to be considered. Recently many media organisations voiced their concerns about keeping churn rates at a reasonable

level after achieving substantial growth. The churn rate describes the ratio of users who cancelled their subscription compared to the total number of subscribers. Because it can be quite expensive for media organisations to win subscribers back, this creates additional incentives to put more efforts in engagement and retention.

3.2. Dynamic Paywalls

A dynamic, adaptive or 'smart' paywall is a relatively new technology that hasn't been defined in literature yet. In this paper, a dynamic paywall is considered to be a technology that at some point in the acquisition process incorporates data-driven personalisation in order to improve its efficiency. Whereas regular paywalls rely on a pre-established set of rules, dynamic paywalls bend those rules individually for every user. For example, users with a high likelihood to subscribe hit the registration or payment prompt earlier than someone who hasn't had enough exposure to the content and the value proposition of a brands services yet. At the core of this paywall scheme is the idea that personalisation significantly increases the efficiency of sales measures and only improves further due to the use of machine learning, which enables powerful computers to dynamically extrapolate algorithms and data models from the growing datasets.

Academic literature on this topic is quite rare, but An et al. described in their 2018 paper the development of an 'Adaptive Paywall Mechanism for Digital News Media', a paywall framework for the Canadian newspaper The Globe and Mail, which based its decisions on the utility and costs of an article as well as the navigational graph (the historic and expected path of articles a user follows during a session). Based on these inputs, the paywall model dynamically decides if the requested article is presented to the user or if a payment prompt will be shown. To train and evaluate the data model, the researchers used a dataset consisting of more than two billion hits, collected between January 2014 and July 2017. By developing and comparing several different approaches, the researchers were able to show that it's possible to significantly improve the efficiency of a paywall with relatively small investments (An et al., 2018, p. 205-214).

But despite this academic proof industry adoption of dynamic paywalls is still quite scarce. Many media organisations announced that they want to develop similar technologies, but failed to deliver so far. Only a handful of media organisations actually use technologies that would qualify as dynamic paywalls, most of them were interviewed for this paper.

3.2.1. Who uses them

The pioneer

It wouldn't be the first time for the Wall Street Journal to be a 'first mover', some even consider them to be the creator of the modern news paywall. The Wall Street Journal began to charge for access to their digital content in 1997, by 1998 they had 200,000 subscribers, in 2007 they reached the milestone of one million subscribers (Cnet, 1998; MacMillan, 2007). In 2014, when subscriber growth began stalling, they decided to change their approach. Because current paywalls put the focus on the content, they decided to adopt a dynamic paywall, that puts the user into the spotlight and adapts to their individual behaviour. Based on this data it calculates a so-called propensity score for every user, which indicates how likely they are to subscribe to the Wall Street Journal. Users with a high propensity score get more access to articles, whereas users with a low score hit the paywall

faster. The marketing efforts also are personalised based on the propensity score. The dynamic paywall considers advertising inventory, so it'll individually loosen to create more advertising capacity or tighten to sell more subscriptions.

The hybrids

Schibsted media group, which owns some of the biggest media organisations in Norway and Sweden, use a semi-dynamic paywall. It is based on their so-called hybrid paywall, a combination of the freemium and metered model, which is still in use at their Norwegian newspaper Aftenposten. Users are able to access a specific number of articles for free (Metered), but some articles were only available for subscribers (Freemium). Svenska Dagbladet evolved this hybrid paywall by adding a fourth layer that is dynamic. An algorithm tries to identify 'engaging content', which will then be put behind the paywall by the newsroom. Despite its success, the dynamic layer doesn't personalise the experience for single users, although they're currently working on a fully dynamic paywall. But they already calculate a propensity score, which is used to personalise targeting for acquisition measures by phone and Facebook ads. The Schibsted media group currently have 1.1 million subscribers in total, 65 percent of those are digital only.

The individualists

The NZZ is currently the oldest newspaper in Switzerland that is still in circulation (founded in 1780) and the most successful broadsheet newspaper of the country. It has had metered paywall for its online platform since 2012, which has been refined in the following years. In 2017 it deployed a dynamic paywall – NZZ calls the technology 'dynamic paygate' – which adapts the meter threshold dynamically and rule-based to the user. Users are able to read up to five articles per month for free on a single device, then they have to register to access up to five more articles for free. Similar to the Wall Street Journal, users with a low propensity score get access to more articles (the maximum amount is still capped at five articles), high-scored users quickly reach a lower threshold. Only registered users are scored, because only then they can be reliably tracked among different devices. With the help of machine learning, data scientists identify new patterns that help to personalise the experience of the paywall, for example by individualising the messaging of the payment prompt and creating exemptions for specific times based on common behaviour, like commuting.

The hopefuls

New York Media announced in November 2018 that the group is going to adopt a dynamic paywall as well. Due to the fact that their subscriptions cover access to all of their brands, such as New York Magazine, Vulture and Grub Street, they want to watch the behaviour of users among all of these platforms and include this in the data model. A user who accesses multiple New York Media brands a month will therefore be confronted with a payment prompt earlier than someone who only reads the same number of articles from one brand. There's currently no information on the effectiveness of their data models, but according to New York Media they have been developed within four to five months and are constantly monitored (Willens, 2018a).

Hearst Newspapers is also experimenting with dynamic paywalls, although currently only at selected outlets. At the local newspaper Albany Times Union, they categorise users in three different segments based on their propensity score. Users with a high propensity score get access to up to five

articles for free and further three free articles after they sign up, users with a low propensity score immediately have to sign up. The data models and scores for users are recalibrated at the beginning of each month (INMA, 2017). Hearst also combines their dynamic paywall with a data-driven targeting strategy. By combining audience and content data, they were able to identify the interests of their loyal audience and could specifically target those segments, like sports fans, with personalised marketing materials (Willens, 2018b).

The New York Times, although still operating with a general meter limit, announced on a recent earnings call, that they'll „be more dynamic about how we use the meter" (Seeking Alpha, 2018). But currently they're only experimenting with different segments, trying to find out if there's a significant difference in efficiency of different price points and meter limits for some segments (Moses, 2018b).

3.2.2. How they work

Almost every dynamic paywall relies on a certain set of rules, that are often based on the paywall model that was in place before. For example, the NZZ still uses the rules of its metered model as baseline and simply made the threshold dynamic. Similarly, the Wall Street Journal used its Freemium model (although most of their content was already locked to non-subscribers, with only a few exemptions) as a basis for the development of their dynamic paywall. The only outlier is Schibsted media group, which, for now, uses the dynamic model only to identify which content should be locked. But because this selection process currently isn't personalised, it's considered to be a supportive hybrid-model, even though it's already technically capable of identifying users with a high subscription likelihood.

Personalisation of acquisition

The dynamic paywall should, at least in theory, improve the acquisition process the most, though its underlying idea - personalisation - can be applied to other parts of the subscriber business as well. By personalising access to content, the conversion rate can be drastically improved compared to paywalls that limit access based on the same rules for their whole audience.

The personalisation is supposed to enable a more targeted approach, where only those who are more likely to subscribe will get content as an incentive to do so. According to the WSJ, this seems to work. The open rate of the paywall is now the same as it was before the introduction of the dynamic model, yet the conversion rate improved by "double percent figures". The open rate signifies how many times content has been accessed without hitting a paywall. This is in line with reports from NZZ, who were able to double their conversion rates with the help of propensity scoring. But according to Barmettler, the benefits from data analysis and translating those into business rules were way bigger. In total, NZZ was able to improve their conversion rates by a factor of five within three years.

Even Schibsted, who only used the propensity score in experiments, were able to benefit from it. They were also able to improve their telemarketing and Facebook marketing conversions by specifically targeting high propensity users. Conversions on the phone improved by six to nine times, the average cost for acquiring a subscriber on Facebook was reduced by 35 percent (Corcoran, 2018).

Details matter

The propensity model can also be used to personalise the experience around the subscription funnel. Someone who hasn't visited the Wall Street Journal before, a 'cold audience', will be greeted with "more brand-led" advertising, whereas someone who has shown a high interest in subscribing, a 'hot audience,' will be offered marketing materials with specific price points, which matters more to them.

Schibsted uses a similar approach with their Touchpoints platform. Touchpoints enables business and product teams to quickly produce and test in-app messages such as banners. They are, after articles, the second most important source for subscription sales. By testing the wording, placement and targeting of these messages, they were able to significantly improve click-through rates and re-subscription rates, but also engagement among current subscribers, for example by reminding them of family sharing and redirecting delivery of the newspaper during holidays (Schibsted Media, 2018). The NZZ also uses personalisation for their marketing materials, where even small details, such as the font colour or the way the pricing is broken down (per week, month or year) are adapted to optimise conversion rates.

Finding the model

In order to reliably assess the subscription likelihood, usually reflected by a decimal number between zero (not likely to subscribe) and one (highly likely to subscribe), organisations first need to identify what kind of data would give an indication of this behaviour. The NZZ evaluated „between 3000 to 4000 combinations of different signals", among them „everything you could think of", says Markus Barmettler, Head of Analytics & Market Research at NZZ. Only those who had the strongest correlation to users becoming subscribers were actually taken into consideration for the development of the propensity model. The NZZ didn't specify how many did end up in the working model, but considering it is subject to constant change, the number of included signals varies as well.

The Wall Street Journal makes use of about 65 different signals, though not all of them are required to calculate a score. Schibsteds propensity score uses ten to fifteen different signals, according to Eivind Fiskerud, currently head of audience development at Bergens Tidende and the former head of data and analytics at Schibsted Norway. Schibsteds model isn't calculated dynamically in real time, but rather every week based on data from the past two weeks. Due to the fact that this data is only used to specifically target high-scored users with telemarketing or Facebook ads, this is sufficient for their needs. The NZZ score is currently only calculated daily – every morning, according to Barmettler – but they're working on reducing this in the future.

Demographic signals, for example relating to education or income, have a strong correlation to subscription likelihood, but due to technical or legal limitations it's often not possible to collect data like this for the propensity model. The Wall Street Journal circumvents those limitations by approximating this data based on other signals, such as the location. There are currently "over a hundred different decision trees" built into the model that try to approximate those signals as accurate as possible, according to Karl Wells, GM of WSJ Membership.

But according to NZZ, demographics weren't as relevant as signals based on the user behaviour and content. „The top behavioural predictors were as we expected, like how many devices are used for access or preferred referrals - direct users have a higher affinity than those who come from news

aggregators like Google News“, says Barmettler. Schibsted also uses a broad variety of signals for their propensity score, though most of them are focused on the behaviour of the user, such as the frequency of their visits, number of articles read, the number of devices used and their respective operating systems.

Search for the needle in the haystack

All organisations interviewed for this paper regularly recalibrate their models based on performance and adapt the weighting or inclusion of certain signals. Often, it’s not just about a simple linear correlation, for example if metric X increases, the subscription likelihood decreases, but machine-learning-supported pattern detection. For example, at NZZ, if a user in general reads only a few articles, but on a variety of topics, they are more likely to subscribe than someone who reads the same number of articles on a specific topic.

Due to the fact that propensity scores are based on probabilities, they’re prone to certain error rates, though no organisation was willing to share specific numbers. But one general rule of thumb seems to be: Not all data is created equal. Therefore, more data doesn’t necessarily improve the accuracy of the scoring model. “It’s not the volume that makes it powerful, it’s understanding what’s in the model that’s most closely related to whether you’ll subscribe or not“, says Wells. Understanding why a certain signal affects the subscription likelihood can substantially help assessing its importance and thus simplify weighting.

Exemptions of the rule

The different models also have certain rules that create exemptions in order to improve their efficiency. The Wall Street Journal created ‘sacred times’ during which historically most subscriptions have been sold and thus demand is probably driven by reasons that currently can’t be abstracted from data, and applies stricter rules during those times. Similarly, the NZZ lifts the paywall during specific times they know a lot of regular users are typically accessing the website, for example while commuting, in order to not disrupt their usual behaviour.

Those exemptions are sometimes identified with the help of machine learning, but often still discovered and translated into rules by humans within the organisation. Schibsted brand Svenska Dagbladet involves sales, editorial and data analytics teams in an constant conversation about the model, though they currently only use it to identify ‘engaging content’ to lock behind the paywall: „We have a stand-up meeting two times a week and discuss if this method identifies the right articles or not and then we adjust it, make exceptions or discuss why it's not working“, says Petter Lorentzon, head of digital sales, analytics and data at the Swedish newspaper.

Content selection

Due to the personalisation of the paywall, editorial teams at Wall Street Journal and NZZ can’t decide anymore if a certain article should be put behind the paywall or stay free. Although this creates a certain friction within the organisation, because the newsroom appears to lose parts of their editorial authority, it is, at least according to those who use dynamic paywalls, the most efficient way. “We’ve avoided a content led model because we did not see a way of attracting and converting people while having our best pieces behind a wall”, says Barmettler. “We need our top content in front of the people.” According to Wells, humans, when given the choice, often make

false assumptions which content should be put behind a paywall. “With the move to customer-led paywalls we found that content types that convert have been different to what we assumed would convert when we let humans choose.”

But there doesn’t seem to be a general consensus on this. News organisations such as Schibsted and Die Welt Online still leave it up to their newsroom to identify the content that should be put behind a paywall, though they provide high levels of data to support the decision making. „We're not trying to replace anything, we're trying to support gut-feeling“, says Lorentzon, who wants to avoid „black-box solutions“ and „autopilots“. At Welt Online, the newsroom uses a so-called Plus-Score that explains with a simple number how well content gains new subscribers and helps retain existing ones. „ Based on this data, the newsroom has learned in the past years, how to pick good content“, says Sven Scheffler, Chief Product Officer at Axel Springer.

3.2.3. Regulation and other obstacles

Despite its potential, the dynamic paywall has to face the same challenges as many other data-driven technologies. Those vary by their nature, some are technical (adblockers and privacy plug-ins, tracking of users among multiple devices), economical (high investment costs, high competition with other industries for skilled employees) or societal („techlash“, loss of trust in technology that requires data collection). But one legal challenge worried most organisations interviewed for this paper: GDPR.

The privacy issues

The European Union introduced the General Data Protection Regulation, GDPR in short, in April 2016. Since May 2018 organisations who are active within the European Union have to abide by this regulation, which enforces stronger data protection and privacy measures, for example by requiring explicit consent for data collection and processing. This made it significantly harder to collect data – the fuel dynamic paywalls run on.

One way many organisations try to bypass this issue is by getting long-term consent from their users by getting them to sign-up for a user account. When signing up, users usually have to, in order to successfully create an account, give their consent to the privacy policy of the organisation, which explains what kind of data is collected and how it will be used. NZZ and Schibsted make us of this, whereas the WSJ technology doesn’t require log-in.

The NZZ originally didn’t introduce its log-in requirement due to GDPR, but to better fight abuse of the meter limit, for example by accessing the website with multiple devices or deleting the locally stored cookies, that kept track of the number of read articles. „By building the model around the user ID, we bypassed the cookie metering.“ Users are able to access five articles for free, then they have to sign up for a free account to read up to five articles more. Although the registration prompt increased the number of logged-in users significantly, „on average still about 50 percent of users aren’t logged in and they still amount for a significant part of our ad impressions“, says Barmettler.

Losing your advantage

Although Schibsted Media has a big network of different online platforms, among them mostly media and classified ad brands, the current legislation makes it harder to use that to their

advantage. „If we want to follow you from Aftenposten to VG, you need to consent to that and if we want to follow you from Aftenposten to Facebook, we need a consent to that as well“, says Bård Skaar Viken, head of consumer business at Schibsted Media Norway. Schibsted owns Finn.no, the biggest online marketplace in Norway, which uses the same log-in-system as other Schibsted platforms. Due to that, most Schibsted media organisations in Norway had quite high log-in-rates and were able to track their users more accurately. This auto-log-in on other platforms has been revoked because of GDPR, which forced them to put more effort into signing up users themselves.

Though Viken doesn't think that log-ins are required to accurately track and score a user, he thinks a lot of valuable data is lost without it: „If we're not able to know how you interact with our content on Facebook or how you interact with your content on other Schibsted Services, it will make it more difficult.“

Lorentzon also said implementing GDPR systems could in fact help with the development of dynamic paywalls in the future: „The GDPR and the structure we had to put in in order to be compliant, helped us to be much more structured when it comes to develop analysis and data models.“

The next big arms race in media

Dynamic paywalls require many expensive investments, among them into technology, skills (by hiring new staff members and training your current employees) and processes. This could lead to a situation where small media organisations with limited resources fail to keep up with the pace set by big brands such as the Wall Street Journal.

But Viken argues that it might be hard for small players to make use of dynamic paywalls anyway. He said it was easier to successfully apply these targeted approaches on national brands such as Aftenposten than local ones like Bergens Tidende. Reaching those „very narrow segments“, for example with Facebook ads, turned out to be very hard to do profitably.

The insights on what kind of data helped the most predicting subscription likelihood also showed that there's probably no „one size fits all“-solution for the industry and every organisation has to adapt to their respective audience and market. Whereas Schibsted is highly focused on different content types and how they can benefit acquisition, engagement and retention, NZZ didn't see any difference in the performance of content types. „We thought that specific content types lead to more conversions, but the opposite turned out to be true. The higher the usage, the higher the likelihood of subscribing“, says Barmettler.

Free dynamic paywalls for everyone

Nevertheless, a common technological framework could help ease the adoption of getting into data-driven processes and paywalls. There are some third-party companies who specifically provide their services and software for the media industry, but those are often too expensive for small media organizations. But there's currently one big exception of this rule.

Slovakian newspaper Dennik N developed and released a machine learning component for their open source platform REMP, which can be used by everyone for free. REMP stands for Readers Engagement and Monetization Platform, which describes an independent set of tools that are supposed to help media organizations monetize their content better. The project is funded with a

grant from the Google Digital News Initiative and according to Dennik N already used by several other media organizations, who offered new ideas for further improvement.

Part of the platform is a machine learning component called Pythia, which assesses the subscription probability for users. Based on the score, media organizations can define rules, for example that only users with a high subscription probability get an offer, whereas users with a low score would get asked to subscribe to the newsletter. The algorithm is built on Dennik N's experiences with conversions and can be applied to most markets without any changes, according to Bella. Because Dennik N usually puts all of their longform content behind the paywall, the algorithm is currently mainly used to personalize marketing measures instead of a paywall meter.

Google could also join the effort to create an industry-wide standard. Googles News Initiative, a program that aims to provide media organizations with resources and technology, announced its 'Subscribe with Google' plan in March 2018. Part of this plan is the development of an 'Propensity to Subscribe' signal, which would indicate to publishers if a subscriber should be shown an ad or a subscription offer. Google announced it was in early stages of testing the signal with selected publishers who used the Google Marketing Platform (formerly known as DoubleClick). Aside from the technology Google offers several incentives to join the 'Subscribe With Google' program, such as higher ranking on search results of active subscribers and an optimized sign-up- and payment-process based on the Google account (Schindler, 2018).

4. Retention and Churn

Most media organisations interviewed for this paper expressed their intention in moving from a focus on acquisition further to engagement and retention. In many ways this seems like a logical step: After building an audience and learning what makes them subscribe, these organisations want to ensure that their existing customers stay with them for as long as possible. In many ways this reflects the principles of a subscription business, which is user-oriented and strives for constant improvement of the customer experience, better than simply trying to acquire more new users than losing existing ones, which is a highly process-oriented approach. But as this chapter shows, the transition from acquisition to engagement and retention is harder than it may seem. From an industry perspective, even the first movers are still just at the beginning of this process.

4.1. Why everyone wants to fight churn

When it comes to the long-term-success of a subscription business, the focus is usually on one key metric: churn. The so-called churn rate is the percentage of users who cancelled their subscription within a certain timeframe. For example, an annual churn rate of twenty percent for a total subscriber base of 100 users would indicate that 20 users cancelled their subscription within a year. For digital media businesses, the churn rate is especially high. According to the current Subscription Economy Index Report by Zuora, media subscription businesses had an annual churn rate of 35 percent in 2018, the highest average churn rate among all industries for this year (Zuora, 2018).

While acquisition metrics such as conversion rates measure success, the churn rate indicates failure. Kjersti Thorneus, director of product management at Schibsted Media Group, recently pointed out that churn is a handicap for every digital business and therefore should be considered to be the most important metric. By reducing that handicap, you vastly improve the effectiveness of your conversion efforts as well. „You don't get money from conversion, it's from retention", Thorneus said (Southern, 2018a).

That's why Schibsted Media, like many other media organisations, shifted their focus from acquisition to retention in order to enable sustainable growth. To achieve that, most organisations currently use two specific measures: churn prevention and dynamic engagement.

Churn prevention is based on the same idea as dynamic paywalls. A data model, often built with the help of machine learning, is using signals that could indicate if a user is about to leave and score them based on that. The higher the score, the higher the likelihood of someone cancelling a subscription. Users who have a high likelihood of leaving are specifically targeted with measures that should create incentives to stay, such as discounts or personalised customer service.

Dynamic Engagement describes a more proactive approach. These kinds of measures use personalisation in order to raise engagement and therefore retention. The higher the engagement, the less likely the subscriber is to churn. Although dynamic engagement has been proven to be a quite effective method, improving engagement with personalisation can be hard and it doesn't guarantee success. Subscribers churn for a variety of reasons, some of them not related in any way to the media group's actions, and therefore it can be hard to influence with automated measures.

The following chapters will show that despite a rise in automation in the area of churn prevention, it still requires a lot of human interaction and mostly involves reasons that can't be influenced by the

newsroom, such as pricing or technical issues. That's why churn metrics are mostly used by business departments, whereas engagement metrics are gaining traction in newsrooms.

4.2. How churn prediction works

Churn prediction is, just like propensity scores, a form of predictive modelling, which uses statistical methods to predict outcomes. The telecommunications and services industry have been at the forefront of this development for more than one decade. But whereas those industries had the advantage of quantifiable data, for example service usage, media organisations often lacked this kind of information and therefore the ability of predicting churn. Only with the rise of digital media and big data did the adoption of those prediction models increase at media organisations as well.

Creating a model to predict churn works similar as it would with propensity scoring. Data scientists try to identify signals and patterns that show the strongest correlation to churn and then build and test models based on that.

Act before it's too late

Though churn prediction is supposed to help raise retention by warning the organisation that subscribers are about to leave, it doesn't necessarily tell them why they're about to leave and if they've already settled for that decision. That's why NZZ abandoned their work on churn prediction models: „The issue with those churn models is that they're trained for the time of renewal. You can only react when a subscription is up for renewal and that's way too late, you need to react when engagement decreases", says Markus Barmettler. The only lever seems to be the price and that turned out too costly for the long term. „We've kind of given up on that for now." That's why the customer retention desk only becomes active after a subscriber already handed in their cancellation request. "This approach has shown to be far more effective than the predictive approach."

Petter Lorentzon from Svenska Dagbladet, who also work on churn models, acknowledges the challenge: „Churn models are a bit tricky, it's not so direct as sales." He still thinks going after churning subscribers is worth the effort, because they've shown interest in the product: „I think, even though they are leaving, it's one of the best leads you can have."

Bård Skaar Viken also disagrees with Barmettler that investing in predictive churn measures isn't worth it in the long term: „I understand that NZZ perspective that actually moving people from one segment to another or from unhappy to happy is expensive and I think this is sort of a broader topic and issue for the media industry, that we are investing a lot of money in marketing and sales, but we are not that willing yet to spend the same amount of money on keeping our customers or sort of retaining them through specific loyalty activities. Because it's expensive and it's hard to see the immediate effect of activities."

4.2.1. Behaviour beats content

As with propensity scoring models, behavioural data seems to show the strongest correlation to churn probability. For the Wall Street Journal the engagement metric that correlates the strongest with churn is "active days", which is the number of days a subscriber has interacted with Wall Street Journal content within a period of 28 days. "The more active you are in a month, the less likely you are to churn", says Wells, who also points out that other metrics are incorporated as well. A low

number of active days doesn't necessarily mean there's a high churn probability for the user, as the model factors in previous behaviour of the individual.

Aftenposten developed a churn prediction model in 2015, which enabled the newspaper to identify customers who are likely to churn and target them with retention measures. They combined multiple data sources to make this possible, such as customer data (demographics and subscription information), CRM data (how and when was the customer in contact with the organisation), digital usage data and third-party data (information like income and education). The data team was able to identify indicators who are strongly linked with churn, such as low digital engagement, manual payment with short renewal periods and days since purchase (Schibsted Norway, 2016). The data is then used to reach out to customers who are at risk of cancelling their subscription in order to keep them onboard. This approach has been extended to the telemarketing and customer support departments of Schibsted media group, which now see churn risk for a customer based on a green, yellow or red indicator (Johannessen & Johnsen, 2018).

Despite using churn prediction and monitoring churn rates, Schibsted Media in Norway also prefers to measure the success of its strategy based on the number of its loyal users instead of the churn rate, as Viken points out. They also analysed their content based on possible correlations to subscription and cancellation likelihood, which led them to create four specific content types users would be actually willing to pay for. Content that is helping the reader understand the news flow (briefings for specific topics), Content that is close to the readers in their daily life (such as financial or crafts advice), Content that is helping the readers understand the world we live in (explainers how certain trends could affect our future as a society) and Content close to readers' interests and identities. Anna Careborg, Head of Premium Content at Svenska Dagbladet, describes the latter as 'Nice-to-know journalism', like restaurant reviews or book tips, which doesn't necessarily lead to conversions, but improves engagement and therefore retention (Careborg, 2018).

Try not to get rid of them in 100 days

Like other organisations, Schibsted is trying to lower churn by raising engagement. „We have the whole customer journey as a playground for data analytics", says Lorentzon. For example, Svenska Dagbladet personalised their onboarding process, moving from time- to user- and trigger-based communication with the Touchpoints platform. Creating habits by pointing out specific features and content types to the user has proven to be a crucial part of churn prevention.

Both the Wall Street Journal and Schibsted are currently focussed on the first 100 days of a subscriber, which seems to be the most crucial time in the lifetime of a subscriber. According to the Wall Street Journal, a subscriber is four times more likely to form a new habit during those 100 days. Habits tend to raise engagement and therefore retention. According to Schibsted, the churn risk substantially decreases (up to 2.5 times) for subscribers who paid more than two bills. Welt Online reported similar results: During the first three months the churn rate was about 50 percent, after that period it dropped to figures of one to two percent.

Forming habits

Similar to Schibsteds approach with the Touchpoints platform the Wall Street Journal created 'Project Habit', where they tried to find out which among the up to 100 ways to interact with their

website have the strongest correlation to retention and how likely users currently make use of them. “You got some things which are really, really sticky, but there’s not many people doing it, which gives you some headroom”, says Wells.

They’ve also analysed content, which showed some correlation to retention. The fact that readers tend to have high loyalty and strong habits when it comes to specific topics and journalists wasn’t new. But the fact that cadence plays a big role in the forming of habits did surprise them. Wells compares the importance of this to programming in TV, to ensure the audience always has something new “they can tune into”.

4.2.2. Search for countermeasures

Knowing who is about to cancel their subscription can be quite helpful in targeting retention measures, but it doesn’t necessarily help finding out why someone is about to leave, which often can be a complex array of different reasons. That’s why it’s hard to actually create countermeasures based on this kind of data, though it helps measure their impact by tracking the change in the predicted churn probability after acting.

Curiously enough doing absolutely nothing can be a quite effective method as well, as it has been noted by multiple people interviewed for this paper. Barmettler from the NZZ agrees with that: „Sometimes it’s smarter to let them rest for a while instead of pointing out that a renewal is coming up.“

Dennik N has a similar stance and tries to improve their pricing strategy, which currently isn’t personalized, with the help of churn prediction data. „If we think the probability of churn is not very high, then we will not offer discounts and we will be careful about not giving a hint that there are some discounts and we would rather wait for another month and then they would probably subscribe again“, says Bella.

For some media organizations, reaching the user can also be a challenge. Someone with a high risk of churning usually already reduced the engagement with the subscription product, so it’s harder to address them with personalized messages. In order to find a solution to this problem, the Wall Street Journal started ‘Project Canvas,’ which Wells describes as a process on how to promote things to a customer that would help them increase their engagement without interfering with the user experience. But he also points out that they’re currently at the beginning of the project, which didn’t surface that many presentable findings yet.

4.2.3. Why humans still matter

Though many media organizations interviewed for this paper hope to reduce churn by automated measures in the future, most of them still rely on telemarketing. Human interaction between user and customer service agent still seems to be the most effective way of getting someone to change their mind about unsubscribing.

The Wall Street Journal even tries to encourage this behaviour by making calls mandatory. Whereas signing up for subscription is possible with a few clicks, unsubscribing is only possible by phone. “I’d say it’s a reflection of the age of our subscriber base”, says Wells, who sees the direct interaction with the customer as a chance to change their mind about cancelling. “People often don’t realize

what they get access to. They don't know that in a week's time we will be rolling out for the midterms, but we tell them."

They also intensified their investments into customer service and use in-house-staff instead of third-party-contractors. Those "brand ambassadors" get trained in-house and deliver valuable insights by talking to customers. "If you think about membership, how much more institutional knowledge you can gain having that capability in-house, that's something we're experimenting with."

Schibsted also uses their churn score to contact users who have a high risk of churning on the phone. The customer service team also has access to the churn score in their CRM system. „If the customer agent can see that the customer calling is most probably on the brink of cancelling the subscription, he can give some targeted communication or an offer", says Viken. Belgian newspapers NRC and De Telegraaf, which are owned by Mediahuis, even try to match the personality of the customer service agent to the subscriber, in order to increase the likelihood of retention (The Newsplexer Project et al., 2018, p. 14).

4.3. Dynamic engagement

Though churn prediction can help identify users who are at risk of leaving, it doesn't actually help fix the underlying issues that lead to unsatisfied subscribers. But considering that most churn prediction models are based on behavioural data, many media organisations believe they've found a lever in engagement. By increasing individual engagement with the help of personalisation, churn probability decreases. This kind of dynamic engagement involves mostly automated solutions, for example by personalising the content selection, timing of an interaction or onboarding processes.

A recent report by the International News Media Association, or INMA, points out that the media industry so far has put more effort onto acquisition measures and predict that the current trend indicates a strong shift towards engagement. The Financial Times for example spends three times more on engagement than acquisition (The Newsplexer Project et al., 2018, p.10). Some organisations like the NZZ already see engagement as the basis of their current growth. „The basis of our dynamic paywall is high engagement", says Martin Jungfer, Conversion and Agile Manager at NZZ. „The more users we engage, the more will see the paygate and this will lead to more success."

Becoming proactive

Bård Skaar Viken from Schibsted Media also believes dynamic engagement could help improve areas which are often overlooked, such as customer service: „We're still very much reacting to our customers and their problems. We are not proactively looking at ways to help a customer and with the technology and the data we have available, it should be possible for us." Currently issues are mostly addressed when they're brought up by the customers themselves, who, according to Viken, are already on the brink of leaving when they do so.

A similar issue is efficient use of channels and platforms, where many organisations show a lack of restraint: „We still communicate way too much to our customers. I think communicating less and actually helping our customers with their basic needs, which is to find the content they want to read or engage with, and to help them find the products or product features that are the right for them." Dynamic Engagement by using targeting can actually help improve the efficiency of this kind of communication while reducing the amount of communication required.

4.3.1. Best Practice

Media organisations try to personalise every aspect of the customer journey in order to raise engagement, which should lead to improvements in acquisition (for new users) and retention (for current subscribers). This involves automating and personalising processes, developing new solutions that encourage engagement and optimising content with the help of data with a focus on engagement.

Personalised communication

NZZ tries to personalise the onboarding process by dynamically selecting the marketing material, content and channel chosen for communication with the user, for example by E-Mail, App or Website. „We calculate the next best action and the preferred channel for every user individually“, says Barmettler. The only part that's currently not personalised is the time the message is sent, though they're planning to do so in future.

For the long term, NZZ wants to personalise the whole customer journey. „We've established within the organisation that we only use pre-determined target audiences and rule-based setups when no model exists or when it doesn't make sense to build one.“ Nevertheless, Barmettler says they created a rule-based matrix (who, when, how, where and what) to decide which parts should be automated with the help of machine learning and which should be rule-based. For example: An E-Mail is sent to new registered users (rule for who) at 10 AM every morning (rule for when) to promote a certain set of articles (personalised with machine learning model for what). „You don't have to hit everything with the machine learning hammer“.

Gamification

In October 2018, the Financial Times launched their their 'Knowledge Builder' tool, which is supposed to improve engagement by applying gamification mechanics to articles. Subscribers can pick topics on which they want to gain more knowledge, for example Artificial Intelligence or Blockchain. By reading further articles on these topics, users can collect 'knowledge points', which are shown on the top of an articles. When the user has read the article, the knowledge points get assigned to their account, which is shown at the end of an article as a progress bar (Scott, 2018).

The knowledge points are dynamically assigned to an article by an algorithm, which scores them based on their content. The tool has only been tested with 13 percent of FT's subscribers, but the results have been promising, especially in raising engagement among subscribers with a high churn risk (Southern, 2018b).

The personalised front-page

Svenska Dagbladet tries to use internal data to automate curation and improve engagement. Their journalists rate every article with a 'news value', ranging from one to five, and a „lifetime value“ - short, medium or long - as well as the selected tags and categories which will be used by an algorithm to find the perfect placement on the front page and other sections of the website. After its placement, the performance of the article - like pageviews, engaged time and conversions - decides if the article will stay in place or be moved (Rodrigues, 2017).

Aftenposten has a similar solution in place, but goes one step further. Only the first three articles are the same for every user, the rest of the front page is personalised based on the user behaviour and interests. Aside from an improved workflow this has led to an increase of ten percent in click-through rates from the front page and a two percent improvement of conversion rates on the front page (Aftenposten, 2018).

Welt Online also currently develops and trials a similar personalised front-page. The newsroom ranks articles based on their priority, an algorithm fills pre-determined spots on this list personalised to the user. For example, a user with a high propensity to subscribe is shown a bigger amount of Welt-Plus-content, which is only available with a subscription. Subscribers also have a bigger selection of Welt-Plus-content on their front-page. Users with a low propensity to subscribe rather get a low amount of Welt-Plus-content in order to create more engagement.

Automated content

Another example of dynamic engagement can be automatic content creation. Although this doesn't seem like a dynamic or even personalised approach, it can be, because it helps create niche content that doesn't necessarily deliver much engagement on average, but due to its volume adds a significant number to the bottom line. MittMedia, which owns multiple local news outlets in Sweden, developed software which turned public data, such as match reports or land registry changes, automatically into short articles. Though they didn't create that much traffic on average - the 25.000 articles created about 2,5 million pageviews within one year - their conversion rates were good enough - they helped create 700 subscriptions in the first year alone - to make the investment worthwhile (Karlsson, 2019).

Leaky by choice

When it comes to manual content selection, there seems to be one main rule: If it's behind the paywall, it stays behind the paywall. All three organisations interviewed for this paper who currently don't use dynamic paywalls usually keep their regular daily news free, whereas exclusive content gets put behind the paywall. Slovakian newspaper Dennik N, whose reporting focusses on domestic politics, differs here. They closely monitor the sales performance of their content. When a specific article doesn't sell any new subscriptions, which usually happens after a few days, access only requires signing up for a free account, which creates new opportunities for engaging a possible subscriber.

„We can't collect any more payments, but we will collect more email addresses for the article", says Bella, who also emphasises, that they try to „squeeze all the subscriptions we can from the article". The articles will also be promoted on different channels, like social media or newsletters, to point out to readers that the article can now be read by registered users. „This was the biggest generator of email addresses for us in the first three years", says Bella. But because resources are limited, those are only selected on a weekly basis. They are also limited to articles which show a "high potential for starting their second life on social media", the majority of their paid content stays behind the paywall. Currently Dennik N has 800,000 unique users per month, 240,000 users have registered with their email addresses and 33,000 of those have an active subscription.

Currently the biggest generator of email addresses and thus leads for Dennik N is the content sharing feature, which is also used by The Information and De Correspondent. Top tier subscribers can unlock paid articles themselves by creating personalised links for sharing. Users who click on those links only need to provide their email address to get access to the article. To further entice the user, the message is personalised, telling them the name of the subscriber who unlocked free access for this article.

A similar social approach is the One-plus-One-offer. Current subscribers can gift a subscription for one month every month to a new user. After the subscription ends, users can renew it for two more months at a discounted price of one Euro. After that period, the subscription is renewed at the regular price.

5. Organisation

The rise of digital media allowed media groups to measure the impact of journalistic content and its reach more quickly and accurately. This rapidly changed the market and forced media organisations to change their structure and make their workflow more dynamic in order to stay competitive. It also required completely new roles within the organisation, for example for data analysis and strategic audience development, and required even experienced employees to learn completely new skills.

Managing this kind of change isn't easy, but it's possible. Most media organisations interviewed for this paper are legacy media organisations who faced dire situations and successfully managed to get themselves out of that, often with the help of data-driven measures.

Basic requirements

One major requirement are standardised technology stacks among all involved departments. This has proven to be a major hurdle for big organisations that house multiple media brands. Schibsted for example standardised their Single-Sign-On-solution, so a Log-in created on Aftonbladet could be used on Svenska Dagbladet as well. But despite that, every Schibsted media brand works with different tech stacks, which makes the development and adoption of new technologies like dynamic paywalls more difficult. „But to fix this we are in the midst of a major project called User Revenue Suite, where our Consumer Business Solutions team is migrating all of our Brands to the same consumer tech platform", said Bård Skaar Viken.

In order to work with data, so-called data scientists and data analysts, who are able to collect, process and interpret the data, are required as well. At the Wall Street Journal, those data experts aren't part of a specific department, but rather embedded within the organisation. Giving data scientists direct access to the organisation and flexibility, enables them to find new insights themselves, according to Wells.

Most media organisations interviewed for this paper created dedicated departments for data analysis and audience development, which closely worked with the newsroom in order to find solutions everyone would benefit from. Both approaches seem to work, though they need to be unified by common business goals, which are translated into actionable KPIs.

5.1. How to find the right metrics and KPIs

Metrics and KPIs are a quite controversial topic. In order to work with data, you need to be able to efficiently collect, process and interpret data. But questions like who gets access to which data, what kind of insights are generated from it and who gets to act on them have revealed deep rifts within several media organisations. By adding a data layer, issues like privacy, workers' rights and business culture were suddenly on the table. Several media organisations who had to face those debates during their transformation said that they appreciated the experience, despite being challenging, because it helped the organisation grow and adapt organically. This also required them to pick their metrics and KPIs meticulously in order to keep their employees involved.

Focus on the winners

The NZZ embraced openness and transparency early on and gave every journalist access to their analytics tool Chartbeat, which mainly looks at engaged time, recirculation and page views. Specific reports, which categorise articles based on metrics as over- or underperforming, are only available to the heads of specific departments, who get to decide how to act on that data. The economics desk for example, realised it was futile covering quarterly results of mid-sized companies, and gave up doing so.

Despite first successes, Jungfer sees a lot of work ahead: „The way the data is handled strongly depends on the individual. We're not really structured yet and still at the beginning." Currently most employees are still getting used to the metrics and try to find ways to incorporate them into their workflows. „If one of our top five most accessed articles has a recirculation of less than ten percent, our newsroom takes action and tries to optimise the related articles, distribution or placement", says Jungfer. „We have a strong focus on the winners, to make the already fast running horse a bit quicker."

According to Jungfer there was barely any resistance from within the newsroom. But some nevertheless refused to accept the validity of the results. „In Switzerland, we tend to polite ignorance. If we don't like something, there's always alternatives available." Some try to defend the relevance of their coverage with alternative reports, for example based on the E-Paper-readership. „That's the curse of digital data collection, that should be somehow unified even more."

Create a demand for learning

A core idea seems to be that the metrics should help create a sense of „we're all in this together", as is confirmed by Bård Skaar Viken from Schibsted Media: „It's not top managements responsibility alone to get our newspapers to actually survive through the digital transformation. That's a mission which is relevant for all the people working here."

Slovakian daily newspaper Dennik N provides all of their journalists, who also own 49 percent of the organisation, access to a self-built analytics tool, that resembles Chartbeat. It's highly customised to the needs of the newsroom, which is mainly focused on the conversion rate. This data is used by the newsroom to optimise the placement, teaser length and titles of articles.

Reporters also get a daily email briefing, that consists of conversion rates, total subscription sales and revenues per article. Numbers like page views, unique users or average time spent on article are never shared with the newsroom. „We just never give them the numbers, like they should not care about, we are not in the advertisement business. But they should care about how many people think that their article is so great that they actually want to buy a subscription", says Bella.

Despite the business model, this didn't affect editorial decisions negatively according to Bella. „The reason why this works actually is that the correlation between what our readers want to pay for and what we ourselves want to write is actually very high." The most successful articles in terms of sales are actually their most acclaimed investigative pieces, which won several awards.

Hunting for the loyal users

Most organisations categorise their audiences by their behaviour. Whereas 'passers-by' and casual readers generally have a lower priority on the agenda of these organisations, so-called 'loyal users' are the ones who everyone is hunting for. Those who regularly consume content with a higher intensity usually have the highest likelihood to buy a subscription as well.

Schibsted Media specifically makes a distinction between loyal users and loyal subscribers. Whereas a loyal user visits the Schibsted brand for a certain number of days per month - this varies among the different brands - a loyal subscriber is someone who paid at least two bills and kept his subscription for more than 100 days. „Engagement is the main reason for people paying the bill in the end", says Viken. „We can see in our numbers that if we manage to get you to stay for 100 days and pay, we see that the churn rate drop."

Hunting metrics

The NZZ is currently trying to refocus their metrics on their subscription business. „An article with a high total engaged time maybe has no touchpoints with subscribers or registered users at all", says Jungfer. „We're currently testing conversion success metrics, which specifically tell us if a user registered after this article and if an article was a touchpoint on their way to a subscription. We want to find out if the article itself was the trigger for the subscription or only a part of the journey towards the subscription."

In terms of acquisition, the NZZ wants to create new metrics that help the newsroom identify how well certain content performs with specific audiences in order to attract new users to their platform. „We need to lure those users onto our platforms with content they don't necessarily expect from us, in order to get them to register and hopefully subscribe later", says Jungfer. While the first generation of those conversion metrics is focused on how well content performs on other platforms, for example social media, news aggregators and search engines, the second generation specifically tries to identify content that gets registered users to subscribe.

Uncertain loyalty metrics

„The third generation is going to help us avoid losing our current subscribers", says Jungfer, who also admits that they currently don't have this kind of data. „We don't know yet if a certain article would help us reduce the churn risk for customers."

Viken points out that reliable engagement and loyalty metrics are hard to act on, as they lag in real time and do not break down into a single number or data point. Compared to acquisition metrics, for example conversions, they need longer periods of time to accurately provide an indicator. „I think it's possible, if you drill down to a pretty granular level, to identify data points that both journalists and commercial teams can look at and act on, but it's hard."

Though the industry is mostly focused on behaviour, Jungfer believes it's important to measure the content as well. „A good product that effectively encourages me to share can't be successful if the content isn't worth sharing. Vice versa it's the same situation if you have great content, but the product doesn't create any incentive to return or interact."

5.2. Metric literacy

In order to work with data, it needs to be presented in a way that is understandable to those who are supposed to act on it. Though this might sound easy in theory, experience tells a different story. Data often gets interpreted in the wrong way - sometimes on purpose, other times because of a lack of knowledge on the technical details - which can lead to bad decisions and have far-reaching consequences on the interpretation of future data. Therefore, a certain sense of 'metric literacy' needs to be established early on.

Stay away from Powerpoint metrics

Svenska Dagbladet tries to raise awareness around metrics and KPIs by communicating the goals they're associated with instead of „something powerpoint-ish that you can't relate to". „It's important that the data and analysis team isn't all about tech", says Lorentzon. Something that is echoed by Viken: „If engagement ends up being something that the commercial teams are working on and not the newsrooms, I think we are more or less screwed."

Data analytics and journalists try to identify what kinds of content perform good and create hypotheses on how to improve their success together. "As long as we don't have a blackbox approach, most of the people are actually very keen to understand and they are very thankful for the information on how to do it better", says Lorentzon, who attributes "90 percent to the process and 10 percent to the data". "It's easy for us to point out the wrong thing, but we would like to be able to do it before we've done it."

Get them involved

According to Bella, the success of KPIs doesn't just require a technical understanding of the metrics, but also understanding how they're connected to the goals of the organisation. „We tried to actually introduce KPIs like the ones we're using right now at our old paper and it was very painful and it was very long and it would have been another 10 years until we would actually get to this point now."

But considering that the Dennik N newsroom owns 49 percent of the organisation and that 80 percent of revenues come from subscriptions, reporters had a stronger involvement in the success of the business and gave the metrics the benefit of the doubt. After a short learning phase during which they saw the benefits of the metrics, they accepted them. „They're writing less and less articles and each article is getting longer and longer because they realised that actually the long form is what sells and the articles that I spent more time on actually sell."

Be annoying

Getting employees to understand is only achieved by actually giving them a chance to learn, according to Martin Jungfer. „When I started at NZZ, I introduced a weekly consultation hour. I also did a monthly KPI show and created several Slack channels for those who were interested." Those who refuse to participate in the beginning will follow eventually, especially if the change happens top down. The executive level gave their commitment by using the data from the audience team for their planning and budgeting, which forced employees to get involved. „The colleagues didn't have a chance to avoid those topics."

The change within the newsroom at Aftenposten was also driven by selected key employees, who were made ambassadors, assigned with promoting the ideas and goals of the project. The role of the editor-in-chief has proven to be especially valuable as change agent, because he has the authority to follow up on the individual performance of journalists, but other well-connected and -liked employees also had great impact in this role. They also tend to celebrate the achievement of certain milestones, which makes working on them more worthwhile. „In the end it's about making change less scary", says Viken.

Make it as easy understandable as possible

Journalists at the Wall Street Journal don't get access to the raw data or specific business metrics. But according to Wells the newsroom will receive briefings curated by the customer insights team, which helps them understand who their audience is and should help boost acquisition. "We don't say: Hey newsroom, here's a fancy score you can use. What's much more useful to the newsroom is knowing who are the people who are buying the Wall Street Journal or are interested in buying it."

Even though other publishers have created more complex metrics that try to translate the vast amount of user metrics in an easy-to-read score, the Wall Street Journal currently refrains from that and focuses for their engagement strategy on the "active days" metric.

Welt Online uses simple scores to make metrics easier understandable for the newsroom. Their article score uses six different metrics which are translated into a score to make its relative performance understandable. They've recently expanded their scoring system with a Plus score, which helps understand metrics important for their subscription business. According to Sven Scheffler, Chief Product Officer at Axel Springer Digital, the introduction of the scores didn't create any trouble within the newsroom. „It rather led to the situation that they felt more in control than before and weren't acting almost blind. The newsroom actually asks for data at a faster pace, so they can adapt more quickly."

In order to make business goals understandable and actionable for every department at NZZ, audience development breaks them down among every department of the organisation. „We try to identify correlations between total engaged time and subscriptions sold per day", says Jungfer. „That should help us create a goal of a certain engaged time per day that needs to be achieved in order to get a certain amount of subscriptions. Everyone gets that."

6. Conclusion

As this paper shows, modern media organisations are part of a digital market that heavily relies on data collection and analysis, which drives paradigms shifts like the move towards user centred design. Producing and distributing content without measuring its most important data points is like flying without an altimeter and compass - it can be done, but getting to your destination is quite risky and would require sheer luck.

Therefore, it's almost impossible finding a modern media organisation that doesn't rely in some way on data in order to improve parts of their business. Many media organisations interviewed for this paper are currently in the process of improving their use of metrics and data, by identifying opportunities to automate tasks with the help of data. This is often supported by more advanced metrics, which are created in a way that employees can act on and understand how they as individuals contribute to business goals.

By interviewing nine key people at six important media organisations who use data to improve their subscription strategy I have been able to answer the following research questions:

- *How do media organizations use data to improve the efficiency of their subscription strategy?* Media organisations with paywalls are increasingly trying to identify users who would be willing to pay for subscriptions with the help of data. Those kinds of dynamic paywalls reflect an industry-wide trend towards behaviour-based prediction models, which not only help identify users interested in subscriptions, but also issues with certain features and content as well as subscribers who are at the risk of leaving. Those technologies, though effective, currently face regulatory challenges, especially in Europe, and require big investments, but there are a few initiatives trying to lower the barrier to entry.

- *Who gets access to this data in what form?* Currently data at most media organisations is mostly available in processed form as metrics or KPIs, which are then provided to the departments they're relevant to. Media organisations who started working with dynamic models also started collecting their unified data in specific data warehouses and data lakes, which are then used by experts, like data scientists, in order to create new insights. Though third-party-solutions like Chartbeat and Google Analytics are still widespread, many media organisations start to opt for self-developed simple dashboards with highly customised metrics.

- *What are the main differences between working with automated and manual paywalls?* Media organisations with manual paywalls heavily rely on data, though not necessarily when choosing which content should be put behind the paywall. Data is more often used as an indicator for the newsroom which content should be further pursued in order to sell more subscriptions or keep current subscribers engaged. The selection of paid content is mostly driven by other simple rules, which are based on historical data and are regularly reviewed. Dynamic Paywalls automate this process in order to make conversions as efficient as possible. This also gives the newsroom the peace of mind to leave this task to the data and focus on producing content.

Though in theory dynamic paywalls would be technically possible at every media organisation that offers paid content, there are signs that indicate this probably won't happen. Especially the fact that data models need to be adapted to the individual media organisation and require constant attention by skilled employees like data scientists - which are currently high in demand and therefore expensive - is a big hurdle for many small and medium sized organisations. But dynamic paywalls can be regarded as a part of the industry-wide move toward automatization and personalisation, which should also help media organisations who don't offer paid content.

Even though technology can be helpful in achieving business goals, its impact shouldn't be overstated. Like in every other business, it's all about a product worth paying for. In media organisations the product is the journalism that is produced in the newsroom. „Nothing beats the content really", says Tomáš Bella from Dennik N. „When there are slow days, when we don't have the kind of Pulitzer Prize thing, then we are playing around with the software and everything, but in the end it's less important than the articles."

Appendix: Interviewees

Name	Organization	Function	Method
Karl Wells	Wall Street Journal	General Manager Membership, Subscription Sales and Marketing	Phone
Elisabeth Kanzi	Spiegel Online	Product Manager	Phone
Sven Scheffler	Axel Springer Digital	Chief Product Officer	Phone
Timo Schürmann	Axel Springer Digital	Director Data Strategy	Phone
Petter Lorentzon	Svenska Dagbladet	Director of Digital Sales, Analysis and Data	Skype
Bård Skaar Viken	Schibsted Media	Director Consumer Business	Skype
Markus Barmettler	NZZ	Head of Analytics and Market Research	Phone
Martin Jungfer	NZZ	Conversion and Agile Manager	Phone
Eivind Fiskerud	Bergens Tidende	Editor for Audience Development	Skype
Tomáš Bella	Dennik N	Director for Online	Skype

References

Aftenposten. (2018). INMA: Best Practice - Front Page Reinvented: The algorithm that supports our journalistic mission - and creates a smarter, more engaging and informative product. Retrieved 26 November 2018, from https://inma.org/practice-detail.cfm?zyear=2018&id=6039FE0B-37B7-46C3-A8EB-99B2D9710F2E

An, A., Davoudi, H., Edall, G., & Zihayat, M. (2018). Adaptive Paywall Mechanism for Digital News Media. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 205–214). New York, NY, USA: ACM. https://doi.org/10.1145/3219819.3219892

Brown, C. (2018, February 27). Helping Local News Publishers Develop Digital Subscriptions. Retrieved 7 February 2019, from https://www.facebook.com/facebookmedia/blog/helping-local-news-publishers-develop-digital-subscriptions

Careborg, A. (2018, November 8). Schibsted Future Report: How to get paid. Retrieved 25 November 2018, from https://futurereport.schibsted.com/how-to-get-paid/

Cnet. (1998, April 15). WSJ reaches member milestone. Retrieved 4 November 2018, from https://www.cnet.com/news/wsj-reaches-member-milestone/

Corcoran, L. (2018, February 12). Not all news site visitors are created equal. Schibsted is trying to predict the ones who will pay up. Retrieved 26 November 2018, from http://www.niemanlab.org/2018/02/not-all-news-site-visitors-are-created-equal-schibsted-is-trying-to-predict-the-ones-who-will-pay-up/

Cornia, A., Nielsen, R. K., Sehl, A., & Simon, F. (2017). Pay Models in European News. Retrieved 20 October 2018, from http://www.digitalnewsreport.org/publications/2017/pay-models-european-news/

Doctorow, C. (2011, March 17). New York Times paywall: wishful thinking or just crazy? Retrieved 19 October 2018, from https://boingboing.net/2011/03/17/new-york-times-paywa.html

Fletcher, R. (2017). Paying for News. Retrieved 7 February 2019, from http://www.digitalnewsreport.org/survey/2017/paying-for-news-2017/

INMA. (2017). INMA: Best Practice - Times Union Dynamic Meter and Paywall. Retrieved 10 December 2018, from https://www.inma.org/practice-detail.cfm?zyear=2018&id=B54D874B-5E40-4415-950F-7DE015F32CEA

Johannessen, S. H., & Johnsen, H. W. (2018, March 11). Schibsted audience analysis indicates churn likelihood, dictates communication. Retrieved 26 November 2018, from https://www.inma.org/blogs/marketing/post.cfm/schibsted-audience-analysis-indicates-churn-likelihood-dictates-communication

Karlsson, S. (n.d.). MittMedia generates subscription conversions with robot journalism. Retrieved 6 February 2019, from https://www.inma.org/blogs/ideas/post.cfm/mittmedia-generates-subscription-conversions-with-robot-journalism

MacMillan, R. (2007, November 5). WSJ.com hits 1 million subscribers. Reuters. Retrieved from https://www.reuters.com/article/us-dowjones-wsj/wsj-com-hits-1-million-subscribers-idUSTON51582320071105

Meedia. (2016, September 12). Relaunch bei Welt.de: entschlacktes Layout, neues Bezahl-Modell und schnellere Ladezeiten › Meedia. Retrieved 19 October 2018, from https://meedia.de/2016/09/12/relaunch-bei-welt-de-entschlacktes-layout-neues-bezahl-modell-und-schnellere-ladezeiten/

Moses, L. (2018a, February 20). Google turns on the charm with publishers by focusing on subscriptions. Retrieved 7 February 2019, from https://digiday.com/media/google-turns-charm-publishers-pivoting-subscriptions/

Moses, L. (2018b, November 8). To goose subscriber growth, The New York Times plans to try a flexible meter. Retrieved 10 December 2018, from https://digiday.com/media/looking-for-the-next-tier-of-subscribers-the-ny-times-plans-to-try-a-flexible-meter/

Nicholls, T., Shabbir, N., & Nielsen, R. K. (2016). Digital-Born News Media in Europe. Reuters Institute for the Study of Journalism. Retrieved from https://ora.ox.ac.uk/objects/uuid:0037eea3-8f97-4bf3-a336-f86e21abfd0b

Pérez-Peña, R. (2010, January 20). New York Times to Charge Nonsubscribers For Unlimited Use of Its Site. The New York Times. Retrieved from https://www.nytimes.com/2010/01/21/business/media/21times.html

Piechota, G. (2018, November 18). What paywall myths and the tale of two Posts say about reader revenue. Retrieved 20 October 2018, from https://www.inma.org/blogs/readers-first/post.cfm/what-if-freemium-is-less-effective-than-executives-think

Rodrigues, F. (2017, March 21). Meet the Swedish newspaper editor who put an algorithm in charge of his homepage. Retrieved 26 November 2018, from http://www.storybench.org/meet-swedish-newspaper-editor-put-algorithm-charge-homepage/

Rozvar, C. (2011, March 17). Everybody's Freaking Out About the Times Pay Wall. Retrieved 19 October 2018, from http://nymag.com/intelligencer/2011/03/everybodys_freaking_out_about.html

Schibsted Media. (2018). INMA: Best Practice - Touchpoints. Retrieved 26 November 2018, from https://www.inma.org/practice-detail.cfm?zyear=2018&id=3C0BDA05-35A6-455D-9FE3-AD407E8BC180

Schibsted Norway. (2016). INMA: Best Practice - Churn Prediction Model. Retrieved 26 November 2018, from https://www.inma.org/practice-detail.cfm?zyear=2016&id=BABC0039-C45E-428B-AA9D9E4D4E307104

Schindler, P. (2018, March 20). The Google News Initiative: Building a stronger future for news. Retrieved 4 February 2019, from https://www.blog.google/outreach-initiatives/google-news-initiative/announcing-google-news-initiative/

Scott, C. (2018, October 17). Financial Times launches new tool to help 'knowledge-hungry' subscribers track their reading | Media news. Retrieved 10 December 2018, from https://www.journalism.co.uk/news/financial-times-launches-new-tool-to-help-knowledge-hungry-subscribers-track-their-reading/s2/a729089/

Seeking Alpha. (2018, November 1). The New York Times (NYT) Q3 2018 Results - Earnings Call Transcript. Retrieved 10 December 2018, from https://seekingalpha.com/article/4217282-new-york-times-nyt-q3-2018-results-earnings-call-transcript

Southern, L. (2018a, November 2). Readers drive 40 percent of Schibsted's revenue. Retrieved 25 November 2018, from https://digiday.com/media/readers-drive-40-percent-schibsted-revenue/

Southern, L. (2018b, November 28). The Financial Times is using game mechanics to lower subscriber churn. Retrieved 10 December 2018, from https://digiday.com/media/financial-times-subscriber-churn-game-mechanics/

Spangler, T. (2018, December 5). Media Companies Trail in Artificial-Intelligence Adoption Amid Fear, Unfamiliarity: PwC Research. Retrieved 7 February 2019, from https://variety.com/2018/digital/news/media-artificial-intelligence-ai-adoption-pwc-study-1203078294/

Sweney, M. (2016, November 3). Telegraph axes metered paywall and launches premium subscription service. The Guardian. Retrieved from https://www.theguardian.com/media/2016/nov/03/telegraph-paywall-premium-subscriptions

The Newsplexer Projects, Whitehead, R., & Wilkinson, E. J. (2018). The Media Subscriptions Blueprint (p. 101). Presented at the INMA Subscriptions Summit, London: INMA.

Willens, M. (2018a, March 5). How Hearst Newspapers changes its paywall to drive reader loyalty. Retrieved 10 December 2018, from https://digiday.com/media/hearst-newspapers-changes-paywall-drive-reader-loyalty/

Willens, M. (2018b, November 13). From hard paywalls to meters to dynamic paywalls: Why New York Media is taking a flexible approach to subscriptions. Retrieved 10 December 2018, from https://digiday.com/media/new-york-media-paywall-subscriptions-flexible/

Williams, A. T. (2016, February 29). How digital subscriptions took over the newspaper industry. Retrieved 18 October 2018, from https://www.americanpressinstitute.org/publications/reports/digital-subscription-take-over/

Zuora. (2018, September). Zuora Subscription Economy Index.