

# Market research report: What has become of new entrants in research workflows and scholarly communication?

Yvonne Campfens\*

*Campfens Management, Amsterdam, The Netherlands*

**Abstract.** Innovation can be defined in many different ways, and can originate both from within existing organizations and from new entrants in a sector (e.g. startups). As with innovation, there are different views on and definitions of the term startup. In general, only two out of ten startups are expected to live up to their expectations and grow. Over the years, many names of new entrants in research workflows and scholarly communication have appeared. What has become of these hopeful new entrants and their products, services, and tools? This Fall, market research was conducted and three questions were investigated: 1. Did they still exist (independently) in 2018? 2. If so, how were they funded and how were they doing? 3. If they were acquired by 2018, by whom and when were they taken over?

To systematically create a list of the new entrants, a range of approaches and sources were used. This resulted in a long list, which was shortened to a sample of 120 independent for-profit startups through various filtering exercises. To answer the three research questions, various approaches were used: desk, literature, and web research; qualitative research (mostly interviews and personal communication); and index research (i.e. Crunchbase and TNW Index).

The most striking first conclusion is that a solid 120 independent for-profit startups were found for our sample, indicating noteworthy autonomous innovation.

The sample used was neither complete nor representative of all new entrants in research workflows and scholarly communication, but proved to be an interesting subset to provide answers to the three questions investigated. The conclusions are thus explicitly valid for the sample only.

Another interesting conclusion of this market research is that the vast majority, 92 of the 120 startups in our sample, was found to exist independently in 2018. Less than a quarter (21%) was acquired by another organization by 2018 (two were acquired twice).

For the startups that existed independently in 2018, the second research question investigated *how they were funded* and *how they were doing*. In terms of funding and revenue, impressive funding and revenue numbers were found: ten startups with more than USD 25 million in funding, and seven with more than USD 10 million revenue (29 with more than USD 1 million revenue). It is however important to note that, through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy.

For the 25 startups that were taken over by another organization by 2018, we found that the overall average amount of time between a startup's founding and its first acquisition was five years, in line with startups in the financial sector. A great variety in the activities of the acquired startups was found, as well as in the acquiring parties.

The vast majority (67%) of the 27 acquisitions were carried out by a party that did not come from a scholarly publishing background. The nine (33%) acquisitions made by publishers show a variety of players, with six different actors in total.

The third prominent conclusion of this market research is that the acquired startups were definitely not all taken over by the big STM players.

---

\*E-mail: [yvonne.campfens@campfens.net](mailto:yvonne.campfens@campfens.net).

The market research leaves questions for further research and invites discussion of possible moderators for success, given the debate on current topics in our sector, for example (technical) disruption and the Joint Roadmap for Open Science Tools.

Keywords: AcademicTech, Acquisitions, Funding, Innovation, Market research, MediaTech, Research cycle, Research workflows, Scholarly communication, Startups, TechForMedia

## 1. Introduction

Over the years, many names of new entrants in research workflows and scholarly communication have appeared. These players aim to provide improvements on solutions for existing needs, or address new requirements or unarticulated needs in all areas of the research cycle. Whether in an academic, governmental, or corporate environment, the research cycle covers a researcher's continuum of generating ideas and hypotheses, obtaining funding and approval, conducting research, disseminating results (including formal publication), and overall management (e.g. finding collaborators, organizing work and projects).

The new players that were observed, whether for-profit startups or players otherwise organized and financed, form part of what can be referred to as "MediaTech" or "TechForMedia," depending on their area(s) of application. Media technology is described on Wikipedia [1] as including, among others, digital media technology, electronic media technology, electronic publishing, and media technology university programs. In line with the Wikipedia definition of FinTech [2] (or financial technology), MediaTech is defined in this paper as "new technology and innovation that aims to compete with traditional methods in the delivery of media services," and thus includes "alternatives for traditional approaches to research workflows and scholarly communication." The latter specifically has become known as AcademicTech. Technology for media, or TechForMedia, is defined in this paper specifically as "solutions and services for media companies" (e.g. publishers).

What has become of these hopeful new entrants and their products, services, and tools? This Fall, market research was conducted to investigate various questions in this respect:

- (1) Did they still exist (independently) in 2018?
- (2) If so, how were they funded and how were they doing?
- (3) If acquired by 2018, by whom and when were they taken over?

This paper describes the approach and results of the market research. The underlying data are available from Zenodo [3]. The author welcomes feedback on and discussion of all parts of the market research.

## 2. Background

Innovation can be defined in many different ways. Some basic views are given on Wikipedia [4]:

*Innovation can be simply defined as a "new idea, creative thoughts, new imaginations in form of device or method." However, innovation is often also viewed as the application of better solutions that meet new requirements, unarticulated needs, or existing market needs. ...in economics, management science, and other fields of practice and analysis, innovation is generally considered to be the result of a process that brings together various novel ideas in such a way that they affect society.*

Innovation can originate both from within existing organizations and from new entrants in a sector (e.g. startups). To this end, Van Peteghem and Mohout [5] distinguish *inside-out* and *outside-in* models in their

book *Corporate Venturing*, which has as its main theme “*accelerate growth through collaboration with startups*.” Under outside-in corporate venturing, a wide range of financial, operational, and technical models are described for such collaboration and organized into two main categories: equity-based partnering and non-equity-based partnering models. Bradley and O’Toole [6] from McKinsey&Company describe how potential disruption from outside can be countered by existing organizations in *An Incumbent’s Guide to Digital Disruption*. They explain why navigating disruption can be challenging, but also that organizations that show foresight and a willingness to make bold moves can emerge as winners.

As with innovation, different views and definitions exist for “startup.” Robehmed [7] addresses the question “what is a startup?” in a Forbes blog post, which includes this view:

*“A startup is a company working to solve a problem where the solution is not obvious and success is not guaranteed,” says Neil Blumenthal, cofounder and co-CEO of Warby Parker.*

Jaime Novoa [8] analyzes what types of fundraising vehicles are being used during the different stages of a startup’s life in his Startupxplore blog post “Understanding Differences in Startup Financing Stages.” He clearly describes how a startup’s revenue typically develops over time (including the famous “valley of death”), with three subsequent stages of financing vehicles:

- Startup financing at the early stage, or pre-seed capital: FFF (Fools, Friends & Family), business angels, accelerators.
- Seed capital: business angels, super angels, early stage venture capital firms, crowdfunding, syndicate investing.
- Startup financing at the growth stage: Series A, Series B, Series C, and following rounds.

According to Omar Mohout, over 1 million euro in funding or revenue is generally regarded to be a valid indication of innovation for a startup in Belgium and the Netherlands. In other regions, these thresholds can be different and are generally higher, especially in the USA. Also, in general, only two out of ten startups are expected to live up to their expectations and grow.

### 3. Methodology

#### 3.1. Sample

To systematically create a list of the new entrants in research workflows and scholarly communication, a range of sources were used, listed below. The number of players pulled from each source is indicated, with some duplication between the sources (duplication not indicated below).

- Over 400 tools and innovations in scholarly communication (charting the creation and availability (supply side)) [9]. This database continues to build on the 2015–2016 survey by Bianca Kramer and Jeroen Bosman (both at Utrecht University Library), who are interested in the way information is created, shared, and processed in academia [10]. From this source, 687 players were pulled (Figure 1). From the answers to the closed and open questions in the 2015–2016 survey, 225 players were selected.
- Presenters at STM events [11] from 2010 to 2017. This added 53 players.
- Outsell’s [12] “Companies to Watch,” as per their annual Information Industry Outlook (2009–2017). This source provided 29 players.
- Pitches at APE (Academic Publishing in Europe) Conferences [13] between 2011 and 2017. Here, 44 players were found.

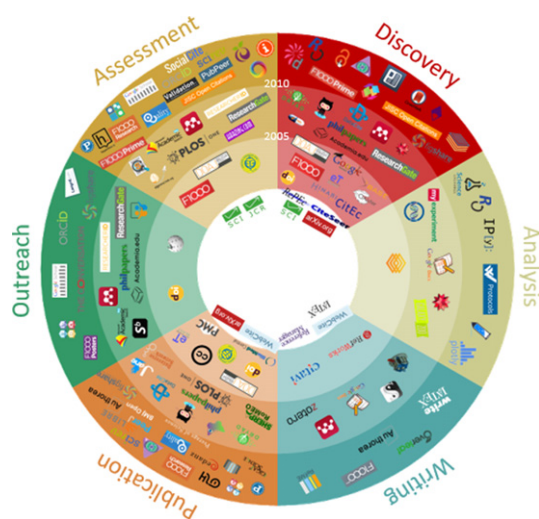


Fig 1. Innovations in scholarly communication and changing research workflows (source: J. Bosman and B. Kramer).

- Nominees for the ALPSP Award for Innovation in Publishing [14]. This source contributed ten players from 2014 to 2017.

In addition, players were found by talking to people and scanning industry news throughout 2018. This approach brought the total long list to over 800 players, which included startups, collaborative projects, launches by established commercial companies, university initiatives, solutions by individuals, EU-funded projects, initiatives subsidized/sponsored by foundations, etc. It is important to note that these 800 players are not all new entrants, nor are they necessarily innovative. The list is to be regarded as a long list of players and their products, services, and tools [15], relevant to research workflows and scholarly communication at large.

In order to bring the long list down to truly new entrants and players relevant to workflows in original research [16] and the subsequent scholarly communication of results, the following filtering criteria were applied:

- Players established between 2007 and 2017.
- Independent players.
- Initiatives relevant to original research and subsequent publication in academic (or scholarly) journals [17].

When applied, this first filter led to the explicit exclusion of inside-out initiatives by established players, which can of course also be an effective approach, as is illustrated by Elsevier's (RELX group's) recent recognition by Forbes as a world leader in innovation [18]. Other examples of well-known names that dropped off the list in this filtering are Bookmetrix and F1000. These were simply beyond the scope of this market research. Also excluded were pharma-sponsored initiatives.

Because the remaining list was still too long to investigate the three questions listed in the introduction above in detail, another systematic approach was applied in order to get to a manageably short list of interesting startups:

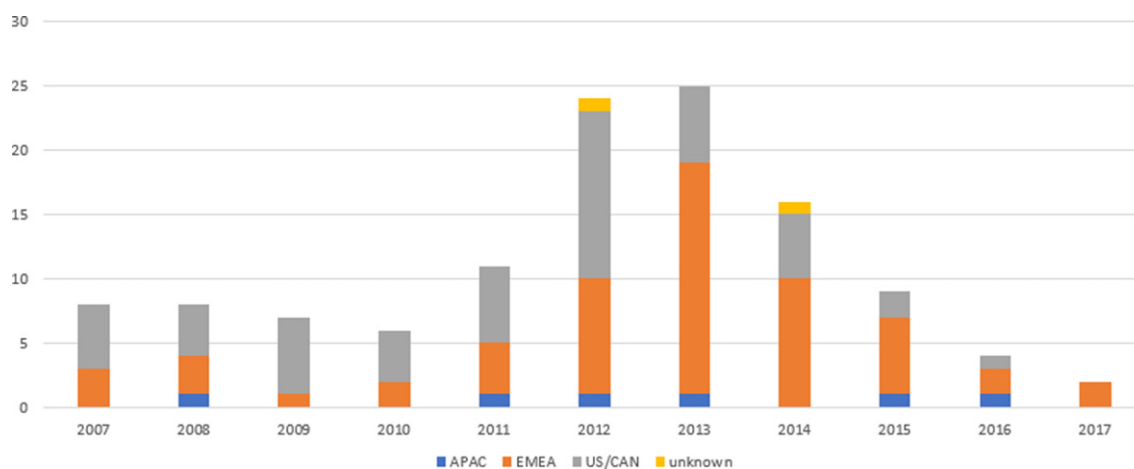


Fig 2. Number of startups in the sample by founding year and region.

- Keep “for-profit” and exclude “not-for-profit” (subsidized, charity, university, EU grants, etc.) initiatives.
- Exclude single developers’ initiatives.
- Cover the first research phases only (as per the above mentioned Utrecht University research) [19]: preparation, discovery, analysis, writing, and publication (“select journal” part only). Thus, exclude part of the publication phase (e.g. “archive/share,” “peer review,” “publish”), and the outreach and assessment phases.

This second filtering exercise saw more initiatives excluded from the list. Some examples are: Coko Foundation, Hypothes.is, Impactstory, Unpaywall, and Zotero. Again, these were simply beyond this market research’s scope.

This approach shortened the list to 120 young, independent, for-profit startups. The sample contained 32 TechForMedia players, often also with a role in MediaTech. The 115 pure MediaTech players cover all the research phases: preparation (21%), discovery (37%), analysis (19%), writing (19%), and publication (“select journal” only, 5%). Often, more than one research phase is covered by a startup.

The founding years of the startups in the sample spread across 2007 to 2017, and show a peak around 2012–2013 (Figure 2). Half of the startups in the sample are based in EMEA; 43% come from US/CAN, and only 5% are based in APAC.

### 3.2. Approach

The market research aimed at answering the following questions for the 120 young, independent, for-profit startups in the sample:

- (1) Did they still exist (independently) in 2018?
- (2) If so, how were they funded and how were they doing?
- (3) If acquired by 2018, by whom and when were they taken over?

To answer the questions, various approaches were used, which were applied more or less interchangeably. The approaches were:

- Desk, literature, and web research.
- Qualitative research (mostly interviews and personal communication).
- Index research (i.e. Crunchbase [20] and TNW (The Next Web) Index [21]).

The most important source for assigning the research phases covered by the startups was the classification used in Utrecht University's database (built on the 2015–2016 survey). This database assigns a single research phase per tool. If the information gathered during the market research gave reason to, relevant research phases were added to the coverage of the startup. In many cases, this resulted in multiple classifications per startup. In a few cases, it was decided to assign a different (single) classification than the one used in the Utrecht University database. It should be noted that the assignment is subjective.

This research was done in the Fall of 2018 and the findings were double-checked by the author at the end of December 2018. It should be noted that the status of the startups in the sample is likely to change in time and that the results of the market research are therefore best regarded as a “2018” snapshot.

## 4. Results

### 4.1. Question 1: Did they still exist (independently) in 2018?

What has become of these 120 promising young startups in 2018? The market research indicates that the vast majority (92 (77%) out of the 120) still exist independently, and less than a quarter (25 (21%) out of the 120) were acquired by another organization. Only three startups in our sample were found to have ceased to exist or, more specifically, two had stalled their activities due to the GDPR (Global Data Protection Regulation).

### 4.2. Question 2: If so, how were they funded and how were they doing?

This market research provided a variety of insights into the 92 startups that existed independently in 2018. First, their geographical breakdown, which differs only slightly from the overall sample: 49% are from EMEA, versus 50% in the overall sample; for US/CAN, this is 46% versus 43%; and for APAC, it is 4% versus 5%.

For 51% of these startups (47 out of the 92), details on their latest funding round are known based on information found in Crunchbase, TNW Index and personal communication. For the vast majority of these 47, this is Seed capital and Series A and Series B (Figure 3). For these 47, we see a slightly different distribution in research phase coverage compared to the full sample: preparation (25% here versus 21% in the overall sample); discovery (30% versus 37%); analysis (25% versus 19%); writing (17% versus 19%); publication (“select journal” only, 3% versus 5%). Of these 47 startups, 17 are also active in TechForMedia.

In terms of parties investing in these startups, we have so far only analyzed strategic investor Digital Science [22]. Digital Science (sometimes in combination with Holtzbrinck Digital) have invested in six of the 92 startups.

For 40 startups (out of the 92), their total funding is known from Crunchbase and TNW Index. For those, we see quite a different geographical breakdown than for the overall sample: 33% are from EMEA, versus 50% in the overall sample; for US/CAN, this is 60% versus 43%; and for APAC, it is 8% versus 5%. The funding of 27 of these 40 startups is over USD 1 million, and ten of these (i.e. Asana, ResearchGate, Doximity, ScienceExchange, Typeform, Import.io, Scientist.com, Yewno, Transcriptic, and BenchLing)

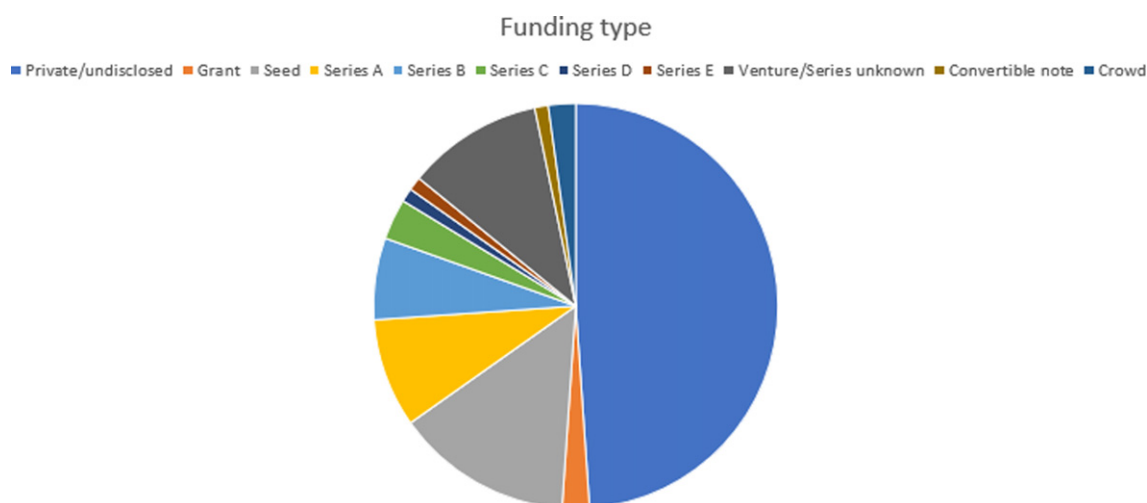


Fig 3. Latest funding round of the 92 independent startups (source: Crunchbase, TNW Index and personal communication).

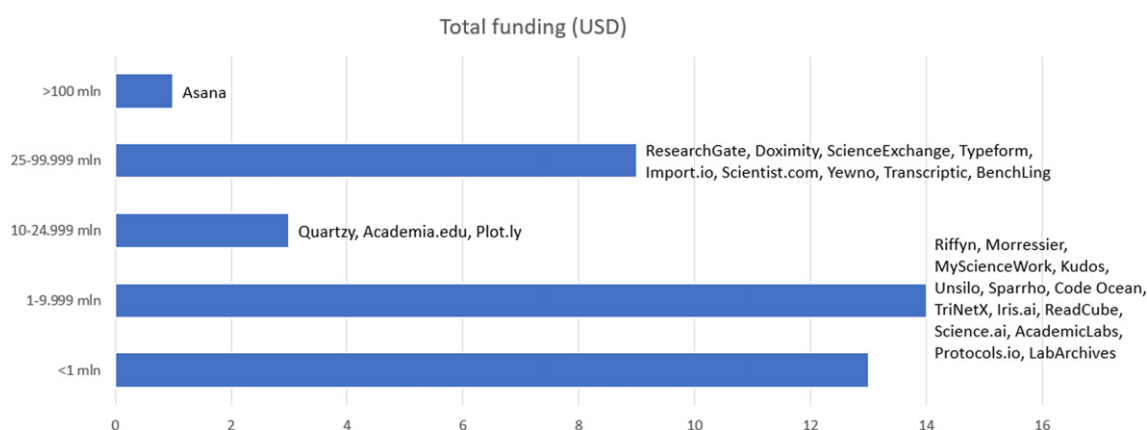


Fig 4. Number of startups with information on total amount invested; the ones with more than USD 1 million funding are named (source: Crunchbase and TNW Index).

have more than USD 25 million in funding (Figure 4). More than half of the top-10 funded startups (six) in our sample cover the preparation research phase. The names of the top-10 funded startups, the research phases they cover, the amount of funding they obtained, and their annual revenue are shown in Table 1.

For two thirds of the 92 startups, their revenue remains undisclosed. Of those whose revenue data are available from Crunchbase (31), almost all have more than USD 1 million in revenue. The revenue of seven (Asana, Scientist.com, ÜberResearch, TriNetX, Doximity, ScienceExchange, Zapnito) is over USD 10 million. In Table 2, the top-10 revenue startups are given in descending order, with information on research phase coverage and funding. More than half of the top-10 funded startups (six) in our sample cover the preparation research phase.

The financial figures in this section are almost fully based on information from Crunchbase. Through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy.

Table 1

Top-10 funded startups (source: Crunchbase). *Note: Through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy*

Startup:	What do they do:	Total funding (USD)	Annual revenue (USD)
Asana	PREPARATION	<b>213,200,000</b>	69,300,000
ResearchGate	PREPARATION/DISCOVERY/ ANALYSIS/PUBLICATION	<b>87,600,000</b>	4,000,000
Doximity	PREPARATION	<b>81,800,000</b>	17,700,000
ScienceExchange	PREPARATION	<b>72,500,000</b>	13,600,000
Typeform	ANALYSIS	<b>52,300,000</b>	2,400,000
Import.io	ANALYSIS	<b>38,300,000</b>	8,100,000
Scientist.com	PREPARATION/DISCOVERY/ ANALYSIS	<b>31,000,000</b>	44,300,000
Yewno	DISCOVERY	<b>30,500,000</b>	n/a
Transcriptic	ANALYSIS	<b>27,800,000</b>	4,900,000
BenchLing	PREPARATION/ANALYSIS	<b>27,400,000</b>	3,000,000

Table 2

Top-10 revenue startups (source: Crunchbase). *Note: Through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy*

Startup:	What do they do:	Total funding (USD)	Annual revenue (USD)
Asana	PREPARATION	213,200,000	<b>69,300,000</b>
Scientist.com	PREPARATION/DISCOVERY/ ANALYSIS	31,000,000	<b>44,300,000</b>
ÜberResearch	PREPARATION	n/a	<b>25,800,000</b>
TriNetX	PREPARATION/ANALYSIS	2,500,000	<b>25,700,000</b>
Doximity	PREPARATION	81,800,000	<b>17,700,000</b>
ScienceExchange	PREPARATION	72,500,000	<b>13,600,000</b>
Zapnito	DISCOVERY	561,000	<b>10,000,000</b>
Import.io	ANALYSIS	38,300,000	<b>8,100,000</b>
Altmetric	DISCOVERY/ASSESSMENT	n/a	<b>7,600,000</b>
ReadCube	DISCOVERY	n/a	<b>6,300,000</b>

In the abovementioned Utrecht University 2015–2016 survey some information is available on take-up and usage of the 92 startups in our sample. Most often mentioned (at least in two categories) are: Academia.edu, ResearchGate, ReadCube, Figshare, Kudos, PaperPile, and Qiqqa. Another 20 players are scored in one category.

#### 4.3. Question 3: If acquired by 2018, by whom and when were they taken over?

Under Question 1, it was indicated that 25 of the 120 startups in our sample were taken over by another organization by 2018. In fact, two of the 25 were acquired twice in the period under investigation (2007–2018). The acquiring parties did not only come from a scholarly publishing background. In fact, the vast majority (67%) did not.



Table 3  
Coverage of acquired startups, industry of acquiring party, and founding and acquisition year details (part 1)

Startup:	What they do:	Acquired by:	About:	From start to acquisition year:													
				07	08	09	10	11	12	13	14	15	16	17	18		
Papers app	DISCOVERY/WRITING / PUBLICATION	Springer (ReadCube)	publisher (workflow tool)	6 years						4 years							
PubGet	DISCOVERY	CCC	IP services	6 years													
Mendeley	PREPARATION/DISCOVERY /WRITING/PUBLICATION	Elsevier	publisher	6 years													
FluidSurveys	ANALYSIS	SurveyMonkey	market research/media	7 years													
Wizfolio	PREPARATION/DISCOVERY /WRITING/PUBLICATION	iGroup	distributor/info services	10 years													
Expernova	DISCOVERY	Questel	IP services	11 years													
Paperrater	WRITING	Barnes & Noble	distributor/info services	10 years													
Scrazzl	PREPARATION/DISCOVERY	CompareNetworks	market research/media	4 years													
Kaggle	PREPARATION	Google	analytics/data	8 years													
Meta (was Scencescape)	DISCOVERY	Chan Zuckerberg Foundation	social/foundation	8 years													

These non-publisher acquirers (the ones with more than one acquisition) are active in the following industries: five (19%) in distributor/information services; three (11%) in analytics/data; two (7%) in biotech; two (7%) in IP services; two (7%) in market research/media; and two (7%) in workflow tools.

The nine (33%) acquisitions carried out by publishers show a variety of players: three were done by Elsevier, two by Atypon/Wiley, one by AAAS, one by Condé Nast, one by Springer, and one by Taylor&Francis.

One fifth (five) of the acquired startups were founded in 2012, and ten out of the 27 acquisitions were done in 2017. The overall average amount of time between a startup's founding and its first acquisition is five years.

The variety in the activities of the acquired startups, the acquiring party, and details about the founding year and year(s) of acquisition are illustrated in the overviews in Tables 3, 4, and 5. In addition to MediaTech, five (20%) of the 25 acquired startups also serve TechForMedia.

## 5. Discussion

This market research looked into what has become of new entrants in research workflows and scholarly communication. In order to obtain a manageable sample to investigate, a long list of players underwent filtering procedures, leaving out a group of potentially interesting initiatives: subsidized, charity, university, EU grants, and single developers' initiatives. Also, players active in the coverage of certain research phases were excluded, namely later parts of the publication phase (e.g. "archive/share," "peer review," "publish"), and the outreach and assessment phases. The filtering also led to the explicit exclusion of

Table 4  
Coverage of acquired startups, industry of acquiring party, and founding and acquisition year details (part 2)

Startup:	What they do:	Acquired by:	About:	From start to acquisition year:													
				07	08	09	10	11	12	13	14	15	16	17	18		
EasyBib	WRITING	Chegg	distributor/info services														
Colwiz (now Wizdom.ai)	PREPARATION/DISCOVERY/WRITING	Taylor & Francis	publisher														
GenomeCompiler	ANALYSIS	Twist Bioscience	biotech														
Plum Analytics	DISCOVERY/ASSESSMENT	EBSCO (Elsevier)	distributor/info services (publisher)														
Poetica	WRITING	Conde Nast	publisher														
ShareLatex	WRITING	Overleaf	workflow tool														
Manuscripts	WRITING	Atypon/Wiley	publisher														
Authorea	WRITING	Atypon/Wiley	publisher														
Sample of Science	DISCOVERY/ANALYSIS	fullstop	publisher services														
HiveBench	ANALYSIS	Elsevier	publisher														

Table 5  
Coverage of acquired startups, industry of acquiring party, and founding and acquisition year details (part 3)

Startup:	What they do:	Acquired by:	About:	From start to acquisition year:													
				07	08	09	10	11	12	13	14	15	16	17	18		
Nowomics	DISCOVERY	Healx	biotech														
Publons	DISCOVERY	Clarivate	analytics/data														
Peer Reviewer Evaluation (PRE)	PREPARATION/ASSESSMENT	AAAS	publisher														
REFME	DISCOVERY/WRITING	Chegg	distributor/info services														
Kopernio	DISCOVERY	Clarivate	analytics/data														

inside-out initiatives by established players. The remaining sample of 120 young, independent, and for-profit startups seems to be an interesting one. The most striking first observation is that this sample of 120 indicates noteworthy autonomous innovation in itself.

However, the overall sample of 120 startups seems to have an uneven distribution of founding years, with a peak around 2012–2013. This can be the result of certain market developments that were not investigated further in this market research. But there is another important factor to consider: the methodology used to compile the sample. The sample is biased towards startups that are already in a later and operational phase of their life cycle. This can explain the lower numbers after the 2012–2013 peak: with this methodology,

we do not yet see interesting startups that are still in their earlier days. As no (free) comprehensive source of startups was known, startups from earlier founding years that failed were probably already invisible in 2018. This can explain the lower numbers before the 2012–2013 peak. It can also explain the low number of ceased startups found in this research.

Only three startups in our sample we found to have ceased to exist by 2018 or, more specifically, two had stalled their activities due to GDPR. Because of the methodology used, this market research does not provide meaningful insight into ceased players (neither their numbers nor characteristics).

The 115 pure MediaTech players cover all research phases, but there is no reason to conclude that the distribution is representative or that it is not, as no comparisons with other sources and research results were conducted.

It is interesting to see that 27 startups are active in both MediaTech and TechForMedia, and that there are only five pure TechForMedia startups in the sample. This can be an effect of the methodology used to obtain the sample, namely the strong contribution from the Utrecht University database (built on the 2015–2016 survey). It can also be an indication of a lack of innovation in the scholarly publishing sector, or of most of the innovation being done in-house and/or by established players.

When considering the results and the geographical breakdown of the sample in light of developments in research and scholarly communication in APAC, more startups from that region were expected to have been found. This raises interesting questions for further investigation: Where can we source initiatives from APAC? Or are they perhaps happening in different ways?

Despite these considerations, the sample with the set of 92 startups existing independently in 2018 and the 25 that were acquired by 2018 proved to be interesting. However, any conclusion is to be regarded as only relevant to the specific sample.

### 5.1. Startups that existed independently in 2018

With respect to the independent startups, the question was asked: *how were they funded and how they were doing in 2018?* It is interesting to see that much more is known about the funding and revenues of startups from US/CAN than from EMEA. This is expected to be related to regulation and culture.

Given the bias of the sample with respect to founding years, the finding that most funding is in Seed capital and Series A and Series B seems logical.

The high percentage of startups with more than USD 1 million in funding or revenue is very likely to be due to the fact that more is known about US/CAN-based startups, in combination with the fact that amounts are generally higher there than in EMEA. Nevertheless, seeing an absolute number of ten startups with more than USD 25 million in funding and seven with more than USD 10 million revenue is absolutely impressive. Among the top-10 funded startups, five are from the top-10 revenue startups. Among the top-10 revenue startups, five are from the top-10 funded startups.

The financial figures are almost completely based on information from Crunchbase. Through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy.

Due to a lack of sources, little was discovered on customer/researcher take-up and usage, and it is difficult to draw conclusions in this area other than that follow-up research and information would be desirable.

What is worth discussing are links to possible moderators for success, given some hot topics in our sector:

- Are these startups ready to disrupt [23–25]?
- Do they fit the Joint Roadmap for Open Science Tools [26]?
- Do they support diversity and inclusion [27]?
- Do these startups use disruptive technologies [28]?
- Do they allow for modular/interoperable services that can be integrated into other systems?
- Do they offer a cross-publisher (neutral) solution?

More information is also desired with respect to the startups that existed independently in 2018. Some open questions that remain, and that are of interest to further investigation in future (market) research, are the following:

- Who are the investors in these 92 independent startups, and what else do they invest in?
- Are these independent startups showing growth in terms of take-up and usage?
- How are they performing against original investment and revenue goals?
- What is their impact on society and science? Do they create jobs? What about customer satisfaction and brand recognition?
- How are these MediaTech and TechForMedia startups doing compared to benchmarks from other industries?

## 5.2. Startups that were acquired by 2018

The finding of an average of period of five years between founding year and first acquisition seems in line with other industries (in the financial sector in 2018, about half of the startups were between three to ten years old when acquired [29]). That less than a quarter (25 (21%) out of the 120) were acquired by another organization is probably not what people working in research workflows and scholarly communication would have expected. Before this market research, a sentiment was picked up about a small number of traditional big players in scholarly communication “taking it all.”

In light of this, the fact that for 67% of the acquired MediaTech startups the acquiring parties did not even come from a scholarly publishing background is interesting. Could it be that we traditionally have a narrow view of our scholarly communication sector?

The variety in the activities of the acquired startups, the acquiring parties, and details about founding year and year(s) of acquisition are interesting. Again, could it be that we traditionally have a narrow view of our scholarly communication sector?

Moreover, it would be very interesting to know more about the sample of acquired startups. Some open questions that remain, and that are of interest to investigate further in future market research, are the following:

- What was the amount they were acquired for?
- (How) Did the startups survive the acquisition and integration?
- To what extent have the acquired startups become part of the core business of the acquiring party?
- To what extent have they contributed to the revenue/EBITDA of the acquiring party?
- Have they grown in terms of take-up and usage after being acquired?
- How are they performing against original investment and revenue goals?
- Were they acquired because of the underlying technology, product or service, or people?

- Or were they acquired to take a competitor/threat out of the market?
- How are these MediaTech and TechForMedia startups doing compared to benchmarks from other industries?

## 6. Conclusion

### 6.1. Answering the three research questions

The most striking first conclusion is that, even with a lot of filtering from an initial long list of over 800 players and their products, services, and tools relevant for research workflows and scholarly communication at large, a solid 120 independent for-profit startups were found for our sample, indicating noteworthy autonomous innovation.

The methodology used in this market research delivered a sample of for-profit startups that was not complete in terms of including all players founded between 2007 and 2017, and did not cover all research phases. It is also assumed to be unrepresentative in terms of APAC coverage. Despite these considerations, the sample with the set of 92 startups that existed independently in 2018 and the 25 startups that were acquired by 2018 proved to be interesting and provided answers to the three questions investigated. The conclusions are explicitly relevant to the sample only.

With respect to the first question (*did they still exist (independently) in 2018?*), the vast majority of the 120 startups in our sample was found to exist independently in 2018: 92 (77%). Less than a quarter (25 (21%)) was acquired by another organization by 2018.

For the ones that existed independently in 2018, the second research question investigated *how they were funded* and *how they were doing*. In terms of funding and revenue, more is known about US/CAN-based startups than about those from other regions. Top-funded and top-revenue startups mostly cover the preparation research phase.

For 51% (47 out of the 92), details on their latest funding round are known, based on information in Crunchbase, TNW Index and personal communication. For the vast majority of these 47, this is Seed capital and Series A and Series B, which is not unexpected. For 40 startups (out of the 92), their total funding is known. Ten startups with more than USD 25 million in funding were found.

For two thirds of the 92 startups (61), their revenue remains undisclosed. Of those whose revenue information is available from Crunchbase (31), almost all have more than USD 1 million in revenue. The revenue of seven startups is over USD 10 million. It is however important to note that, through private conversations, it is known that many startups do not disclose (revenue) numbers, and that the Crunchbase data are subject to a spread of accuracy.

For the third research question, the 25 startups that were taken over by another organization by 2018 were looked into, and it was investigated *who* did the acquisition and *when*. The overall average time between a startup's founding and its first acquisition was five years, and two of the 25 startups were acquired twice in the period under investigation (2007–2018). A great variety in the activities of the acquired startups was found as well as in the acquiring parties. In addition to MediaTech, of the 25 acquired startups, five (20%) also serve TechForMedia.

The vast majority (67%) of the 27 acquisitions was executed by a party that did not come from a scholarly publishing background. Among them, we found organizations in a range of non-publishing

industries, such as distributor/information services, analytics/data, biotech, IP services, market research/-media, workflow tools, etc. The nine (33%) acquisitions done by publishers show a variety of players, with six different actors in total.

The overall conclusion about the acquired startups is that they are definitely not all taken over by the big STM players.

## 6.2. Questions for further research

This market research provides meaningful answers to and conclusions about a subset of new entrants in research workflows and scholarly communication. It would be interesting to do research on the other subset(s) as well, especially on startups covering the research phases excluded in this sample and the independent not-for-profit players. It would also be very interesting to find out more about new entrants from APAC.

With respect to customer/researcher take-up and usage, little is discovered and few sources are available, so specific follow-up research and information would be desirable in this area.

Moreover, in other areas, more information would be interesting with respect to the startups that existed independently in 2018, as well as the ones that were taken over. This could include, for example, information regarding investors, performance, and comparisons with benchmarks from other industries.

Also, it would be worthwhile to discuss links to possible moderators for success, given the debate on current topics in our sector, for example (technical) disruption and the Joint Roadmap for Open Science Tools.

## Abbreviations

**APAC:** Asia Pacific

**EMEA:** Europe, Middle East & Africa

**GDPR:** General Data Protection Regulation (EU) 2016/679

**TNW:** The Next Web

**US/CAN:** USA & Canada

## Declarations

### *Acknowledgments*

The author appreciates the encouragement from Michiel Kolman and Joris van Rossum to do this market research in the first place, and acknowledges the advice and contributions of ALPSP, Jeroen Bosman, Toby Green, Bianca Kramer, Nivida Lamichhane, Eefke Smit, Ivo Verbeek, Adam Warren, and Harald Wirsching. The input from Alex Ballering, David Sommer, Jörgen Balk, Bettina Görner, Daniel Hook, Joyce Haas, Ivo Hunink, Omar Mohout, and Milan Wielinga when discussing the topic under investigation is also recognized. The author shows gratitude to all mentioned above and to the others who served on the advisory group of this project: Nicolien Adema, Alison Mitchell, and Koji Yamashita.

### *Funding*

The market research covered in this paper was funded by the author herself.

### *Availability of data and materials*

The datasets used and analyzed during the current study are available from:

- <https://zenodo.org/record/2530048#.XC9Ss1xKjIV>
- <https://zenodo.org/record/2538328#.XDtp9VxKjIX>

### *Competing & financial interests*

The author declares that she has no competing interests. The author works as independent consultant at Campfens Management, holding a freelance and participation relationship with Peerwith BV.

### *Open Access*

This white paper is distributed under the terms of the Attribution-NonCommercial 4.0 (CC BY-NC 4.0) International License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. You may not use the material for commercial purposes.

The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

## **References**

- [1] [https://en.wikipedia.org/wiki/Media\\_technology](https://en.wikipedia.org/wiki/Media_technology).
- [2] [https://en.wikipedia.org/wiki/Financial\\_technology](https://en.wikipedia.org/wiki/Financial_technology).
- [3] <https://zenodo.org/record/2530048#.XC9Ss1xKjIV> and <https://zenodo.org/record/2538328#.XDtp9VxKjIX>.
- [4] <https://en.wikipedia.org/wiki/Innovation>.
- [5] Corporate Venturing, 2018, Dado van Peteghem Omar Mohout, ISBN 978 2 87403 494 7.
- [6] <https://www.mckinsey.com/business-functions/strategy-and-corporate-finance/our-insights/an-incumbents-guide-to-digital-disruption>.
- [7] <https://www.forbes.com/sites/natalierobehmed/2013/12/16/what-is-a-startup/#671fee7c4044>.
- [8] <https://startupxplore.com/en/blog/types-startup-investing/>.
- [9] [https://docs.google.com/spreadsheets/d/1KUMSeq\\_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/editgid=0](https://docs.google.com/spreadsheets/d/1KUMSeq_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/editgid=0).
- [10] <https://101innovations.wordpress.com/>.
- [11] <https://www.stm-assoc.org/events/?previous>.
- [12] <https://www.outsellinc.com/>.
- [13] <https://www.ape2019.eu/ape-literature>.
- [14] <https://www.alpsp.org/Awards>.
- [15] <https://zenodo.org/record/2530048#.XC9Ss1xKjIV>.
- [16] <https://en.wikipedia.org/wiki/Research>.
- [17] [https://en.wikipedia.org/wiki/Academic\\_journal](https://en.wikipedia.org/wiki/Academic_journal).
- [18] <https://www.relx.com/our-business/our-stories/relx-group-recognised-as-a-world-leader-in-innovation>.
- [19] [https://docs.google.com/spreadsheets/d/1KUMSeq\\_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/editgid=1519702055](https://docs.google.com/spreadsheets/d/1KUMSeq_Pzp4KveZ7pb5rddcssk1XBTiLHniD0d3nDqo/editgid=1519702055).

- [20] <https://www.crunchbase.com>.
- [21] <https://index.co>.
- [22] Daniel Hook: “Digital Science has the desire to transform the research space in a positive manner helping all the stakeholders (funders, publishers, research organisations [public, private and commercial], governments, and most importantly, researchers) to access modern technologies to support their research portfolio”.
- [23] “The Future for Academic Publishers Lies in Navigating Research, Not Distributing It,” Mattias Björnmalm, January 2018. <http://blogs.lse.ac.uk/impactofsocialsciences/2018/01/29/the-future-for-academic-publishers-lies-in-navigating-research-not-distributing-it/>.
- [24] “Do we need to ‘fail fast’ to achieve open access?” Toby Green, September 2018. <http://blogs.lse.ac.uk/impactofsocialsciences/2018/10/22/do-we-need-to-fail-fast-to-achieve-open-access/>.
- [25] After the IPO was over... “...in the forefront of revolutionising the workflow of the bench scientist and materially increasing the return on the global RD budget...,” David Worlock, May 2018. <http://www.davidworlock.com/2018/05/>.
- [26] <https://jrost.org/>.
- [27] <https://c4disc.org/>.
- [28] <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/disruptive-technologies>.
- [29] <https://www.statista.com/statistics/885893/global-startup-age-acquisition/>.