



## **RELX - STM Investor Seminar**

9 November 2022

Transcript

### Disclaimer

This transcript is derived from a recording of the event. Every possible effort has been made to transcribe accurately. However, neither RELX plc nor BRR Media Limited shall be liable for any inaccuracies, errors, or omissions.

Kumsal Bayazit:

Good afternoon. I'm Kumsal Bayazit, the CEO of Elsevier. I joined RELX in LexisNexis almost 20 years ago. Since then, I've had multiple operational and strategic roles in every part of the business. I have worked in Risk, Legal and Exhibitions divisions, as well as at RELX Corporate as Chief Strategy Officer, and as chair of the RELX Technology Forum. I've been the Chief Executive of Elsevier since early 2019. I'm delighted to be sharing with you today how STM or Elsevier fits into RELX, what it is that we do, and our financial performance. I will cover how we're executing on our strategy to continue to improve our growth trajectory. I will also share how we deliver on our mission of helping researchers and healthcare professionals advance science and improve health outcomes.

In databases, tools and electronic reference, Maxim Khan, Cameron Ross and Josh Schoeller will give you examples of how we serve our customers with high value decision tools in academic and government, corporate and health segments. Jill Lubber, chief technology officer who recently joined Elsevier from risk division will follow and highlight the significant role of technology at Elsevier. I will then come back to wrap up the presentation and do the Q&A session. Let me start by putting in context the STM business within RELX. STM represents about 34% of the revenue and 40% of the profit of the RELX Group.

In the last 12 months to June, 2022, our business generated £2.7 billion in revenue or \$3.6 billion. By format, almost 90% of our revenue is electronic. We are a truly global business with customers in more than 180 countries. About 75% of our revenue is subscription-based primarily in multi-year contracts. Our transactional revenues cover almost the whole product portfolio, and a large part of it is recurring in nature. Let me spend a few minutes talking about our business, what it is that we do with our almost 9,000 team members worldwide.

We are the leading provider of high quality primary research. We have a long track record as a leading STM publisher with titles like The Lancet turning 200 years old next year. We're number one globally by volume and quality with the largest selection of journals for authors, journal brands that we constantly develop as scientific research and disciplines evolve. We have the leading research and analytics platforms in STM publishing. Our database and tools and electronic reference are now nearly 40% of total revenue. What we do here is apply artificial intelligence and sophisticated analytics on the content and data assets to derive

high value insights for our customers, similar to Legal and Risk businesses at RELX, and we do this in a number of segments globally, academic institutions and governments, R&D intensive corporations, health providers and educators.

In each one of these, we have differentiated data and analytics assets with leading positions. We're building comprehensive product suites and expanding use cases. Finally, we continue to provide print such as journals, reference materials and books where there's customer demand. We serve customers of every size. We serve thousands of institutions and millions of individual researchers, R&D and healthcare professionals and students. It's a highly fragmented customer base with no one customer or customer group accounting for more than roughly 1% of revenues. In academic and government segments, we work with virtually all leading research organisations worldwide. In the corporate segment, we focus on R&D heavy industries like life sciences, biopharma and engineering. In health, we have solutions for most stakeholders in the health ecosystem, healthcare providers such as hospitals, pharmacies, payers, as well as medical and nursing schools. In each of these customer segments, our objective is to be a trusted partner to the customers we serve and be known for quality.

We operate in long term global growth industries. R&D spend is critical for nations and corporations to create competitive advantage, drive innovation, economic growth, and solve societal issues such as climate change. As a result of this, R&D spend and number of researchers in both academic and government and corporate segments is large and has grown 4 - 5% annually for decades. As we all live longer lives and aim to live healthier lives, health expenditures and number of physicians and nurses are growing strongly as well. In both R&D and health segments, there's increasing information intensity and exponential growth in datasets. This increases the need for high value decision tools for the customer segments we serve.

We were starting to see an improving revenue growth trajectory driven by our focus on analytics and decision tools in early 2020, just before the pandemic hit. After a brief slowdown, we're now back on that improved trajectory. Underlying revenue growth improved 3% in 2021 and has accelerated further to 4% year to date in 2022. Our objective is to continue on this improving growth trajectory. The majority of our cost base is not directly linked to revenue. A large portion of our cost is technology infrastructure and systems. We have a strong focus on process

innovation and use of artificial intelligence and machine learning to improve our revenue growth, whilst keeping our cost growth below revenue growth. The business has low capital intensity, and capex to sales ratio is very similar to the risk division.

Our improved growth trajectory is driven by the change in our business mix. Print has declined historically in mid to high single digits, and is about 10% of the business, down from just under 25% in 2015. So print drag continues to reduce. Academic and government primary research has grown in low single digits. Our focus here is continued growth driven by article volume growth with high quality research on leading technology platforms at the lowest effective unit cost. Corporate primary research has grown slightly faster in mid-single digits, and we're focusing on expanding datasets and adding more sophisticated analytics to serve specific corporate R&D needs. The main driver of our growth acceleration has come from the increased sophistication in our analytics offering, which you can see here under databases tools, and electronic reference, which has grown in mid to high single digits and is now nearly 40% of our revenue, up from 25% in 2015.

It is this change in mix that is driving the improved revenue growth trajectory for the STM division, as databases tools and electronic reference and corporate primary research, which together are just under half of the divisional revenues, become a larger portion of revenue and print drag decreases.

Now let's have a look at Primary Research first. We serve a lot of different stakeholders here, but at the heart of this process, our researchers and their discoveries. And what we do here is help validate, improve, and disseminate their scientific findings. Researchers get funding from governments, funders and corporations based on their research proposals. They conduct the research at the end of which they have findings. They codify these findings in an article to share their contribution and enable other researchers to build on this knowledge.

Our portfolio of over 2,800 journals help researchers verify and share their findings. We get over 2.6 million submissions per year to our journal portfolio. We attract submissions because of the quality reach and brand recognition of our journals. These author manuscript submissions or pre-prints, as we call them, are reviewed by in-house or external editors. We have publishing teams who constantly recruit

leading experts in their fields to work as editors for our journals. These editors do the first review of the article, and after this stage, about half of the submissions are rejected, either because they're not the right fit for the scope of the journal, or the research is not sound science.

When an article passes the editorial review, these editors are responsible to find two to three experts in the field to peer review the work. Editors find these peer reviewers using their own network, as well as AI driven tools we provide. Editors coordinate the peer review process with our systems. During the peer review, the author will respond to questions from the peer reviewers, improve the paper, and sometimes conduct additional work. At the end of the peer review, roughly half of the papers peer reviewed are accepted for publishing by the editor. Then comes the production stage. It's critical that the paper does not have any mistakes, figures are correct, and images are accurate. We undertake work to edit and ensure that the paper is ready for consumption. During this stage, we also add metadata, include links to optimize the article for online reading.

Then we distribute the article via our platforms, which has significant reach and search engine optimisation and increasingly sophisticated analytics. So, it drives usability and readership, which is really important for the author, and vital for ensuring quality research is shared and built upon by others. Where an article is published, the quality of the journal and its reach is important. It drives citations, collaborations, funding, and ultimately supports researchers career goals and their impact on society through patterns and innovations, policy and public debates. Journals over time accumulate citations by the papers they publish. There are various ways to measure these citations. Journals who are the most highly cited develop reputations for being impactful and high quality over time. We have the largest and have the highest quality portfolio, and scale matters. Barriers to launch an individual title are low, but successfully developing a journal with a trusted community and running at the scale with reach creates competitive differentiation and sustainability.

The main driver of the long term growth in primary research is the strong article volume growth. Over the last decade or so, article submissions to Elsevier have tripled to more than 2.6 million submissions per year due to the breadth and quality of our journal portfolio. There was a spike during Covid period, and as we come out

of the pandemic, we're still seeing increasing submission volumes compared to pre-pandemic. The number of articles published has doubled from 300,000 to 600,000 in the same timeframe. Now, I told you a bit about what we do, how we support the research ecosystem and the services we provide to our authors and customers. Now, I want to spend a couple of minutes talking about the payment models.

Traditionally, this market has operated on a pay to read model where readers or their institutions as users of content pay and authors publish for free. Over time, an alternative model has gained traction where authors or their institutions or funding bodies prefer to pay to publish their research so it is freely available to read. The latter model is commonly referred to as open access. Both pay to read and pay to publish models are available as a subscription or on a transactional basis to our customers. We provide both payment models for our services as well as combinations of the two models to support our customers diverse needs and preferences. In recent years, we have seen a higher rate of growth in the paid to publish model starting from a low base. And you can see in 2021 about 20% of articles published was from pay to publish payment model. Revenue is roughly of a similar scale. There can be a range of lower or higher percentage points depending on how you attribute value. Our combined deals can include several components, pay to read, pay to publish, and database and tools, and are often on a subscription basis. The key point here is that we're here to serve our customers in any way that they would like and we work collaboratively with them and support them to achieve their research goals.

Let me talk about quality as that is a key driver of submissions. A measure of quality in research is looking at citations. How many times an article is cited that's used by other scientists is a proxy for impact. I will show you here a few different ways of measuring that impact and quality of our portfolio. The first graph shows that we have the largest overall share of all articles published at 18% and our share by citation is higher at 28%. Showing research published by Elsevier gets cited ahead of industry average. If you normalize those citations by their field as different disciplines have different average number of citations. For example, chemistry articles on average will have 14 citations and mathematics will have five. We are again, the leader with 1.42 field weight citation impact and the world average is indexed at one. If you look at our portfolio with over 2800 journals, we have the largest selection of journals and we cover all quality tiers as we want to find a home

for every article that is sound science. Today, our portfolio is heavily weighted towards the top tiers as you can see in the graph. It's a virtuous cycle where the quality of our portfolio is attracting high quality submissions as authors want to be published in trusted journals in their discipline.

The other key value driver for researchers is reach. They want to be able to reach as wide an audience as possible for their research to have an impact on progressing their discipline, contributing to other research, to policy, and to innovation. We have industry leading research platforms with true global scale and visibility. The platforms currently attract more than 100 million monthly visits from 50 million unique monthly login ins, and that has grown over 25% gather over the last several years resulting in about 1.8 billion unique downloads this year. The volume of content is large and the quality is excellent, which fuels growth. We use machine learning and artificial intelligence to index the content, drive more efficient search and smart recommendations. We believe that as science becomes increasingly more multi-disciplinary, reproducible, and transparent, and as the volume of research continues to increase and researchers increasingly share their data sets, our platform usability, analytics capabilities, the use of artificial intelligence and machine learning to surface what is relevant and uncover connections and patterns will over time become even more helpful to researchers.

Let me now cover Databases, tools and electronic reference. Databases, tools and electronic reference help our customers solve critical and complex problems. We serve three primary customer segments, academic and government, corporate and health customers though a lot of our solutions are used across customer segments. As I told you earlier, database and tools and electronic reference together with corporate primary research is just under half of total divisional revenues. Of that just under a third comes from academic and government customers just under one third from corporate and a bit over one third from health customers. What we do in each one of these segments is broadly similar. However, the objectives and needs for our customer segments differ. We take our underlying data sets and technology capabilities and build analytics and decision tools that help our customers make the right decisions based on their needs. For example, how to allocate research funding in the academic and government segment, find the right chemical compound for a drug in development and corporate or choose the right oncology pathway for a patient in health segments. We have extensive product suites that serve each of

these segments and we continuously experiment and innovate to support new use cases for our users. Vast majority of our focus is on organic development. This is where we also do small acquisitions of data sets or capabilities where we can accelerate our organic approach.

These higher value analytics and decision tools are underpinned by four key capabilities. We have deep customer and domain knowledge, foundations of which you can find in our primary research business. We employ a lot of researchers, PhDs, doctors, and nurses. We have been serving these customer segments for more than a century in some cases so there is real depth of understanding. We're continuously expanding our data sets. You heard a lot about primary research, but we also collect vast quantities of other contact and data sets, such as research information, patents, policy documents, grants, drug databases, and medical claims data. We apply advanced linking capabilities. We extract entities such as researchers, topics, diseases, genes, symptoms, and add sophisticated analytics, and we do this on global modular scalable platforms leveraging our 2,600 plus technologists and RELX's technology scale and capabilities. Now, to bring this to life, Max, Cameron, and Josh will illustrate how customers in different segments use our tools. As I mentioned, we have multiple product suites for specific use cases. We're going to share a few of them with you today to illustrate how we add value to critical customer decisions with our products. I will hand it over to Max now who will cover academic and government segments.

Maxim Khan:

I'm Maxim Khan senior vice president for analytics products and data platform. I joined RELX in 2007 and over the last 10 years, I've had the privilege of serving our academic and government customers. Globally, governments, funders, and corporations invest around \$2.4 trillion annually on research and development. These stakeholders need to manage the impact of their investment. Our databases and tools help our academic and government customers take high value decisions across the research ecosystem. For example, government funders need to decide where to invest money and track investment outcomes to demonstrate value back to taxpayers. University research offices need to define research strategy, measure performance, and attract funding in order to do best in class research serving their communities. Researchers need to find the best journal to publish in and researchers to collaborate with in order to do research and progress their career. In

general, better outcomes are achieved through collaborations that are interdisciplinary international and those involving industry.

Meanwhile, competition for funding continues to grow with universities looking at evidence-based approaches to compete for funding and do the most impactful research. Our linked data powers our analytics products and services. Our data is differentiated for three reasons. Firstly, it's deep with over 2.8 billion links around 250 million records across multiple data types. Secondly, it's accurate. Our linking accuracy is recognised by our customers to be ahead of competition. Finally, it's authoritative. Our data is used by key influencers behind resource allocation in research. For example, the field weighted citation impact, an indicator built on top of our data is used at national level when looking at research impact. We collect public sources at scale, use machine learning to create links and offer customers flexible options to consume our data. The differentiation of our data means that products based on it are used by many of the research intensive universities across the world.

SciVal is an example of one of our analytics products based on our link data. SciVal enables our customers to analyze research performance, develop partnerships, and understand research trends. We launched it in 2014 and have grown the number of modules to six with revenue growing at around 20% CAGR in the last five years. I'm going to demonstrate several of the SciVal modules by adopting the persona of a director in the research office in a leading UK university. I'm going to use SciVal to inform how we position the university to grow research and funding in strategic areas. First, I'd like to see active areas of research at my institution, which have significant momentum globally and which may be good candidates for funding growth. In SciVal, I can do this by looking at topic clusters within the overview module. The bubbles are topic clusters for my institution computed on Elsevier citation graph.

There are 1,500 topic clusters comprising 96,000 topics. A bigger bubble means more research at my university as measured by publications. Momentum for a topic also called prominence is based on data such as journal quality metrics and the citation graph of over a billion citations. I want to see high momentum topics in which I also have a footprint. These are candidates to get focus from funders, governments, and collaborators. If I look more closely, I can see strengths at the top of the list such as algorithms and computer vision models. As I look further down the

list, I can see topics with significant global momentum where I only have a moderate footprint such as this RNA related topic here. I'm going to see funding for this topic using the grants module and see if there's an opportunity to grow my research and my funding in this space. I can see this topic has received significant competitive funding globally. I can also see institutions who capture funding globally. I can also see institutions who capture funding in this space. From this global view, I can target recruitment where I can drill down into individuals winning the competitive funding and doing the research. As most funders are national and I'm based in the UK, I want to see how I stack up in the UK. I can see from this list that I'm ninth, so there's possibly room to grow. Using the trends' module, I can see the volume of my research output in this topic and it's been flat recently. I can also see the impact of my research in this topic by using an indicator called the Field-Weighted Citation Impact, which measures the impact of research. I saw the volume of research output was flat, and now that I can see that the impact of our research in this topic has also been flat recently. We can look at collaborations to strengthen our research impact and competitive position for winning funding. Using the collaboration module, I can identify existing collaborations in this topic that we may be able to target for expansion.

I can also look at potential collaborators who are active in this topic, who are not yet collaborating with, including those in industry. I can see from this that there could be scope to spin up new collaborations in the U.S. I can actually drill down to individual collaborations, even looking at individuals in those universities and organisations who are doing research on this topic that I could potentially reach out to. In summary, as a research administrator, I can identify areas of my portfolio that have significant momentum globally, drill into their funding and identify ways to increase research impact and share a funding through partnerships. I will now hand it over to Cameron who will talk about analytics and decision products for our corporate research customers.

Cameron Ross:

Hello, I'm Cameron Ross, managing director for Elsevier's Life Sciences Solutions Group. I joined Elsevier in 2004 and today I will share how our team of passionate life sciences and technology experts create world leading data and analytics to help the world's largest pharmaceutical companies make new discoveries, develop targeted medicines and devices, and help ensure compliance and drug safety. Each year, more and more data and insights are produced by both corporate researchers

and their academic peers. In 1950, the estimated doubling time of medical knowledge was 50 years, but just late last year in 2021, medical knowledge was expected to double every 73 days. To make decisions that involve large investments and high risk impacting human health. We bring together extensive scientific, technical and medical content, powerful analytics and advanced technologies to help our customers in pharmaceuticals and biotech, medical technology, chemicals, as well as educators and students in chemistry and health, invent and commercialize ideas, products and processes.

We have a wide suite of products. Today I will share two quick examples with you. First, I'll do a demonstration of our new AI retrosynthesis module on Reaxys, and second, I'll show you how we are expanding into new use cases with our recent acquisition SciBite. We deliver our range of life sciences solutions in a flexible, modular way based on customer needs and use cases. That can be as one of our products, or it could be by delivering our content directly to customers via API in a machine readable format, ready to be used by their data science team. Or finally, as data orchestration solutions, we help our customers extract meaning out of any content. Let me focus first on Reaxys. Reaxys is a well established leading chemistry research platform, underpinned by a comprehensive content set, 265 million substances, 60 million reactions, 100 million documents, and 43 million bio activities, and we are adding new content sets such as an extra 35 million patent records added in 2021 with sophisticated analytics.

One of the key areas in drug development is deciding what molecules to make and how to make them. A research team may need to create 2 to 300 new drug candidates per sub project, design and synthesis planning can take between 8 and 12 weeks per candidates, burning a huge amount of time and effort. 80% of known disease targets are being pursued by multiple competing companies at the same time, so being first to discover a novel approach is critical. We have worked with leading scientists in the field to design a predictive model that uses neural networks on 15 million reactions and 400,000 rules of chemistry extracted from our content to build a predictive resynthesis tool that learns transformation rules from data and learns to prioritise them and predict reactions. The module was launched in 2021 and some of our largest global pharma companies are using it with very positive feedback.

Let me show you a quick example of how it works. This is the homepage for Reaxys, the most innovative chemistry information platform that supports researcher workflows and digital transformation by providing high quality data in combination with advanced analytics and AI. The user interface is simple and intuitive for chemists who normally search by using chemistry specific language or structures. Let's start by searching for a well known target. This is a really hot target for researchers looking for new ways to develop CDK4/6 inhibitors, which could become blockbuster drugs because they interrupt the growth of cancer cells. By digging deeper, I can analyze targeted results from millions of peer review journals and patents where the chemistry entities are extracted by Reaxys. In this example, I'm really interested in the latest cutting edge research that typically appears first in a patent and because of our deep indexing and data structure, I can also limit by competitors or collaborators like this one. Now, this top result is really interesting. It contains many antibody molecules that could be used in combination therapies to treat and stop the advancement of particular cancers.

Now, analysing all the target molecules here, I'm really interested in this one because it's actually from an experimental drug that is yet to be approved. Now that I know what the target molecule is, the next question facing any chemist in a research team is how could I make this in the most effective, efficient, and legally compliant manner? Clicking 'create retrosynthesis plan' will generate viable evidence based roots in minutes, something that could previously take weeks. Now viewing the results will give me two options. The blue results are existing published roots from patents and journals, but I'm more interested in these green predictive roots because this is where I will have freedom to operate by creating novel alternatives. Let's select this one. It's a simple one step route, and the algorithm is a high competent score, it can be made. Finally, using the tree view, it will show even experienced chemists, unknown or unforeseen route they could take, what preparatory materials they could order and the associated costs and any experimental procedures to consider.

This entire process would normally take weeks, but using Reaxys Retrosynthesis module takes minutes, often revealing roots I might not have routinely developed myself. But finally, I can export this to the company's electronic lab notebook, which is integrated with the Reaxys, accelerating the time to move from idea to feasibility experiments with my research team. So, in summary, the Reaxys Predictive

Retrosynthesis solution is tried, trusted, and valued by top pharma and leading academics because the tool gives time savings of up to three times faster than alternatives for designing synthesis roots. And furthermore, it's not black box AI as it shows literature precedence examples which underpin the predictive roots. And in that focused example around chemistry, it is clear that high quality and curated data are indispensable for effective R&D.

But looking broader, pharma companies manage vast amounts of data, perhaps licensed from third parties or even from their own proprietary research. All of these data come in unstructured formats using different terminologies and hundreds of different naming conventions. For example, a single gene or protein may have dozens of different or overlapping names. With SciBite tools we can now extend to new use cases and help customers apply sophisticated search and analytics across all of these different previously silo data sources. Applying the same approach we apply to our own data, turning unstructured data into structured machine readable data, extracting entities such as proteins, genes, diseases, compounds, enriching and normalising data, and then finding links and patterns, for example, which gene is correlated with which disease presentation. By helping customers connect all relevant data together, our clients can access genuine insights faster, helping accelerate new discoveries and innovations. I will now hand over to Josh.

Josh Schoeller:

Hello, my name is Josh Schoeller. I lead two of RELX's healthcare businesses where we focus on providing solutions that drive better health outcomes, create operational efficiencies and lower costs across the healthcare ecosystem. The global healthcare market is large and diverse. In 2022 spend is projected to exceed 10 trillion globally with the U.S. healthcare system accounting for more than 40%. The industry also faces many challenges such as clinician shortages, shifting financial models and patient safety and quality outcome achievement. We are well positioned to leverage deep data, content and analytic assets to help address these challenges and achieve our mission to improve every patient outcome. Our solutions deliver value across the healthcare ecosystem. Our primary customers are health providers, health authorities and payers. Users include physicians, nurses, and other allied health professionals as well as patients. By leveraging our proprietary content as well as data and analytic resources and technology, we help clinicians make more informed decisions, drive optimal care treatments, and enable better health outcomes.

We help educate patients to aid in their own recovery. We work across health institutions to help identify and address risks, ensuring patient safety and regulatory compliance. And we do this across a broad solution portfolio and we continuously drive innovation and enhancing to existing solutions as well as extending into new use cases. Let me share a couple examples from our solution portfolio we just launched earlier this year.

ClinicalPath is our evidence-based decision support and analytics solution for oncologists. We continuously extend coverage and add modules. It now covers more than 95% of medical and radiation oncology and is used at 56 cancer centers across the U.S. by more than 15% of U.S. medical oncologists treating more than 450,000 patients. It is an excellent example of our evolution from evidence-based content to point of care applications.

The main benefit of ClinicalPath is that it makes excellent cancer care more accessible, ensuring that every patient receives the best treatment for their condition and medical profile. It thereby helps reduce the unintended variability of care and supports improved clinical outcomes and lowers the cost of care. The way we do this is by delivering high quality evidence based clinical pathways directly into the clinical workflow. The pathways incorporate the latest clinical studies, which are prioritised and interpreted by 24 oncology expert committees with 385 experts. Pathways are personalised to the patient by incorporating patient specific data points. For example, age, gender, cancer type, cancer stage, biomarkers and mutations and comorbidities. We have recently added a new module, the Clinical Trial Manager. 70% of patients state they'd be interested in a clinical trial. Pharma companies are also interested in driving clinical trial programs which help them advance new treatments.

Despite this being a win-win for all involved, less than 4% of trials reach the target accrual rate. Why is that? The key reason is lack of awareness from both clinicians and patients. ClinicalPath, Clinical Trial Manager utilises artificial intelligence to map local clinical trials to patient presentations and surfaces the matches at the point of care. Upon implementation, our centers have showed a 40% increase in trial accruals. This helps advance science and offer alternatives to care. The other example I wanted to share shows an extension into a new use case.

Gravitas is our recently launched next generation tokenisation technology, which can be used to de-identify and link healthcare data assets to generate complete longitudinal views of a person's health journey. We do longitudinal views of a person's health journey. We do this by leveraging LexID technology from our Risk division.

Let me walk you through a quick example. A large researcher is conducting a study to identify factors that influence breast cancer tumor progression. To do so, they need various data sets, genomic data, mammography images, medical claims, social determinants of health and mortality data. Gravitas enables these siloed data assets to be de-identified at the source and then linked with referential precision to deliver a single integrated research data set. With it, researchers can now answer what biomarkers influence tumor progression, what socioeconomic factors influence mortality, how can we measure tumor progression to the specific drug and treatment regimen, and much, much more.

In closing, there's no shortage of data or content in healthcare, but there is a lack of proper management, contextualisation and transformation into meaningful insights and decisions, and that is what RELX is driving for our customers, as well as creating better health for all.

Now, let me turn it over to Kumsal.

Kumsal Bayazit:

Thank you, Josh. As you have seen, databases tools and electronic reference help our customers solve critical and complex problems across our three key customer segments, focusing on many specific use cases in science and health with multiple product suites. We help high-value decisions by combining our deep customer and domain expertise, leading data and content sets, analytics and technology platforms.

There are two primary ways we drive strong growth in databases and tools. First is by adding new dataset and analytics modules to existing products. Second is through new products for new use cases. Today, you've seen examples of the first with cyber modules and retrosynthesis with Reaxys, and you have seen examples of the second with SciVal, Clinical Pathways, and Gravitas.

You have also seen how we leverage our sister companies' capabilities. For example, Reaxys leveraging the patent data sets from Legal. And Gravitass leveraging Risk technology and applying LexID to solve healthcare use cases. There's constant experimentation and innovation that drives the product pipeline for new high-value use cases.

I will now hand it over to Jill Lubber, our Chief Technology Officer, who joined Elsevier a year ago after almost two decades in the Risk division.

Jill Lubber:

Good afternoon. I am Jill Lubber, Chief Technology Officer at STM. I joined STM in January of 2022, but have been part of the RELX Group for over 19 years. Previously, I was a technology leader in LexisNexis Risk Solutions. During my time there, I worked in the data engineering and entity resolution space where I helped create identity linking technology, which allows us to accurately link billions of data sources.

Technology is a key enabler at STM. We have more than 2,500 technologists worldwide, representing close to 30% of our employees, a similar ratio as RELX overall. We leverage our global scale, not just at STM, but also at RELX. RELX spends 1.6 billion on technology annually, and we are able to leverage these resources, capabilities and infrastructure at STM.

We help build new products and data and technology platforms that fuel all data and analytics solutions, and we ensure these platforms are reliable, scalable, and secure. For example, technology enables product development and primary research to improve the author experience, submissions, peer review and editorial processes. We use machine learning and AI capabilities to improve our database and tools offerings to enhance customer value at scale.

We also leverage technology to improve operational efficiencies. Let me take a moment to give you an overview of the extensive data assets we have at STM, and how we are building a data fabric layer to enable data analytics to improve customer value and support the improving growth trajectory of the STM division. We have over 1.2 billion unique data points and data connections that grow daily. We are moving towards a common data model. Our common data models include entity types such as authors, physicians, drugs, medical symptoms, and organisational

hierarchies. We use natural language processing to extract data from the unstructured text.

It is not dissimilar to what we do at our Risk or Legal divisions. At Risk, there is a more well defined set of entity types, people, organisations, digital devices, locations and cars. Data supporting these entities also comes almost entirely from structured data inputs. We have a similar but more complicated challenge in that we have dozens of entity types and that number continues to grow. The data used to populate and define these entity models is structured and unstructured text. While this makes the problem space more complex, it also makes the solution more valuable to our customers.

Now, let me tell you how we deliver insights and analysis to our customers. Here you can see our technology stack brings together data inputs to deliver high-value analytics and decision tools for our customers. We start with content and data sets. There are billions of records of structured and unstructured content coming into our data processing pipeline.

We do three main things in our data platforms. We extract data, applying our data engineering processes to extract entities. We enrich data, normalising data into existing data models and annotating it. We link data, applying high quality and extensible natural language based entity tagging and machine learning to connect disparate data points at our rate of over 300,000 new links every day.

These data points and concepts can then be classified using our proprietary domain knowledge data sets created by subject matter experts like researchers and physicians. We're then able to take the knowledge graph that results from the data processing engine and quickly and accurately deliver reliable solutions to customers allowing to build decision tools and analytics. We help oncologists recommend the right clinical trial for a patient, a researcher find the right research collaborator for an interdisciplinary field, a chemist find the best drug synthesis approach. All examples Max, Cameron and Josh gave you today are supported by this approach.

We also leverage our technology and artificial intelligence and machine learning on our own internal processes. To give you one example, identifying the right experts who are willing to review an article is a very time consuming task for editors. We peer review more than a million articles with at least two peer reviewers per year.

Using machine learning, we can help editors find multiple qualified peer reviewers in seconds, identifying and ranking the right experts from more than 40 million records.

Finally, let me give you one example of how this works in practice and what value it delivers to customers. We developed our proprietary healthcare knowledge graph five years ago and it now powers our semantic search and a number of other use cases for customers. With millions of relationships and concepts between medical entities like diseases, drugs, procedures and symptoms, we can link any of our own content or third-party content with this knowledge graph.

As an example of how this is helpful at the point of care, we ingest patient data from hospital systems and use the healthcare knowledge graph to classify the data elements found in the records. We can take a patient's medical information, classify the information as symptoms, diagnoses, treatments, current medications, drug adverse reactions, and so on, to support physicians and nurses make the most accurate clinical decision with the right information.

I hope this gives you a sense of the depth, breadth and scale of our data assets at STM and the scale of our technology capabilities. I'll now hand it back to Kumsal.

Kumsal Bayazit:

Thank you, Jill. Before I wrap up, I'd like to share how we partner with the communities we serve to contribute positively to societal progress. This is very much part of the Elsevier DNA and the way we run the business.

We focus on four primary areas. First is access to knowledge. There are many ways in which we do this, but let me give you some examples. We provide access to research during public health crises. Our COVID-19 Information Center was up and running on January 26th, 2020 before the World Health Organization declared COVID a pandemic. We also provide free access to primary research for patients and caregivers, as well as researchers and healthcare professionals in developing countries. We develop free video tutorials to educate medical students, patients, and caregivers on rare diseases.

Second, area of focus is inclusive research and inclusive healthcare. This year was a highlight for us as we launched the first ever three-dimensional, full female anatomy model for medical students and doctors. We train nurses on how to support LGBTQI patients through our simulation products. Lancet Editorial Board is

now 50/50 from a gender balance perspective, and we're making good progress in gender and geographic representation across our editorial boards.

At RELX, across all of our businesses, we focus on supporting and enabling UN sustainable development goals with our information assets and our reach. Some examples at Elsevier are mapping the global research onto the SDGs to assess contribution by different stakeholders. Also, through our Elsevier Foundation, we support healthcare professionals and early career researchers in the Global South.

Research and innovation is also going to be a critical enabler of meeting the net zero goals. We recently published an in-depth net zero report that examined climate research, which makes up 5% of global research output, versus 1% 20 years ago. We also contribute to progress via mechanisms such as the Lancet Countdown on health and climate change with 43 indicators, or new journals such as One Earth with a focus on multidisciplinary research and sustainability.

We will continue to work with the communities we serve to leverage our unique capabilities, which includes our expertise, resources and scale, and through our products and services to support societal progress.

To conclude, the key messages I would like you to take away from today's sessions are, we help our customers solve critical and complex problems, we operate with leading positions in attractive global growth markets, we're delivering higher value decision tools by combining content and data sets with powerful technology and sophisticated analytics, and I am confident that we can continue on our improved growth trajectory through further evolution of our business mix.

This concludes our presentation. Thank you, and we'll now switch over to Q&A.

Call Operator: If you would like to ask a question, please press start one on your telephone keypad. We will take a first question from Adam Berlin with UBS.

Adam Berlin: Hi, good afternoon. Can you hear me okay?

Kumsal Bayazit: Yes I can. Adam, go ahead.

Adam Berlin: Thanks. Good, thanks very much for the presentation. It was really helpful. I just suppose I wanted to focus a couple of questions on the process you described

about article submissions and how they go through the system. I suppose what I'm trying to work out is if you receive 2.6 million submissions a year, is there a world where you could publish a much larger proportion of those articles than you do today? And so kind of your overall thoughts on that. And specifically, how siloed are the editors? So, if someone submits to a specific journal, an article, and the editor says, "This isn't right for this journal," and then rejects it, you said 50% get rejected at that stage, could that have stayed in the Elsevier ecosystem somewhere? And how are you trying to fix that and so that you can publish more articles? And is that rejection rate different for open access journals versus traditional journals? And is that why you're seeing so much growth in open access, because it's easier to publish more? So yeah, that's the kind of what I'm trying to understand. So, thanks very much for helping.

Kumsal Bayazit:

Sure. Thank you, Adam. That's a great question. You are right. We can publish more of the articles that are submitted to us. Today, we reject about 50% of them after the editorial review because they're not the right fit for that particular journal. But presumably, given we have 2,800 journals, they might be the right fit for another journal. And there's various different things we have underway to ensure that we help the authors find the right journal in the first submission. And if they don't find in the first submission, support them with their next submission so that they can actually get accepted into a journal at the shortest time possible, because speed to market for authors is very important as well. So, that's not how editors work historically. So, there's a change management aspect of that, where over time, it gets easier for us to switch journals from one to the other.

We had good success with these when we did that with ecosystems of journals. So, we can put together five or six journals together and the editors work together, and then we actually are able to accept a higher number of the journal articles submitted to us because we can find the right home for them when there's an ecosystem of journals. And to do that in larger scale, we'll take technology and change and different ways of working, and we're making progress on that every day. In terms of your second question, which is rejection rates for open access versus not, so whether it's paid to publisher or paid to read, the publishing process the articles go through are exactly the same, regardless of the payment model. So, the acceptance rates are really driven more by the quality tier of the journal. It's not driven by whether the article is open access or subscription based. So, the higher the quality

of the journal, the lower the acceptance rate is usually, and that's just really based on the quality of the article, not whether it's open access or subscription based. Thank you.

Adam Berlin: Thank you very much.

Call Operator: We will now take the next question from Nick Dempsey with Barclays.

Nick Dempsey: Yeah, good afternoon. I've got three questions, please. So, the first one, on your slide 14, it shows article submission spiked up in 2020. Unsurprisingly as academics couldn't get into their labs, they were spending time writing their research. Also, I guess COVID helped itself, articles about that. Then roughly held that in '21. If I look at that chart and just continued the current trajectory from '16, '18, '19, we wouldn't get the '21. So, should I worry about that falling off in '22, '23? Second question, are the revenues that databases and tools plus electronic reference quite weighted to the US? And if so, are there opportunities to drive growth in that area by selling more of those internationally? And third question, for an outsider to the industry, the time between a researcher submitting an article and seeing it published and made available to the community seems really long. Players like MDPI have been pioneering speeding that up. Are there things you guys can do with technology and efficiencies to make that process much faster? Because you mentioned, so that is very important for authors, and you want to attract more articles.

Kumsal Bayazit: Right. Okay, thank you very much for those questions, Nick. Let me just take them in order. In terms of submissions, you're right, there was a real spike as a result of COVID, and that comes through in the data. But if you look at growth over time, it's been around 10 to 11% over a decade. And what we're seeing this year is that submissions continue to grow from the elevated levels that they were at. So, I would expect similar growth going forward, but obviously COVID had a bit of interruption with initially a lot of research coming out of the researchers, and then maybe it's slower last year because they haven't been able to get back into their labs. The second question you had was databases, tools, and electronic reference, is there more rated to the US? The way I think about it is really more by segment.

I think for academic and government segment, we have quite good penetration across the globe for corporate. For large corporations as well, quite good penetration across the globe, and there's opportunity to grow further. Where it is

weighted to the US more is health, and we are seeing some real opportunity to grow that in international markets as well. And now, your last question was speed to publish. That is an important, important fundamental driver of research, and you really definitely want to balance a very quality editorial process and a very quality peer review with the right peer reviewers with speed to publish as well. But we have a lot of different initiatives in place that enables us to cut a few days out of that speed to publish every single year. And we continue to improve our speed to publish, and we will continue on that trajectory. Thank you very much. Thanks, Nick.

Nick Dempsey: Thank you.

Call Operator: The next question comes from Sami Kassab from BNP Paribas.

Sami Kassab: Good evening Kumsal. It's Sami Kassab from BNP Paribas. I have my three questions as well, please. First, in light of the transition to open access, do you expect your journal subscription revenues to continue to grow, or do you think that eventually they will decline, albeit offset by growth in open access? And if you think subscription revenues can continue to grow, why that, given the amount of content that flows into open access? Secondly, would you be able to comment on the profit margins and monetisation levels in the two business models? You have the same margins in open access then you have in the subscription model. And if not, perhaps when would you expect both to reach parity? And lastly, Pearson has been talking about a turn, a turnaround in a higher education business. Do you think your medical textbook business can also return to positive growth in the medium term? And if not, then why are you keeping these medical assets? Thank you very much.

Kumsal Bayazit: Thank you, Sami Let me just take your questions one at a time. I think your first question was around OA transition and what happens to subs. So, let me tell you how I think about this business. Ultimately, what we're delivering are articles, and those articles, as you have seen, our submissions have grown in low double digits and article growth has been around 7% over the last decade. That is the primary growth driver in our primary research business. There are two different payment models for those articles. It can be pay to publish, which is open access, or it can be pay to read. And what we're seeing is increasingly both of those models, customers want to via a subscription or transactionally. So, I think that as long as our article

volume is growing and we feel confident that it will continue to grow, we should be able to grow our primary research business, and there will be different payment models which will be driven by the preferences of our customer segments, but I don't think that's an impact on growth.

In terms of margins and monetisation between open access and pay to read or pay to publish and pay to read articles, again, as I said, the article goes through the exact same publishing process regardless of the payment model, so the margins are exactly the same for those. What will drive differences is, again, quality tiers, because in our high quality journals we may actually process 95 articles for every five articles we accept, and that would be a lower rate at lower quality tier. So, it doesn't have to do anything with the payment model, but more as the tier of the journal. And then in our higher education business, we have textbooks for both nurses and doctors. And we have really evolved those business over time to be able to provide both continuity from being a student to actually being a practitioner in a clinical setting, so you can actually use all these tools and services as a student and continue in your clinical practice.

And the books are actually quite important assets in that evolution. So, I mentioned, for example, a 3D anatomy model that we have that's a very, very sophisticated digital product that does male and female anatomy instruction. Obviously, the fact that we have Grays Anatomy, and Netters actually drives a lot of differentiation for our 3D anatomy models as well. So, that's the way we think about our books business in general, is actually a continuation from student to being a clinician, as well as from a print electronic to digital to database and tools. Thank you, Sami.

Sami Kassab: Thank you.

Call Operator: And again, as a reminder to ask a question, press star one. We will now take our next question from Matthew Walker with Credit Suisse.

Matthew Walker: Thanks a lot, everyone. Yeah, I've got three questions, please. The first is you mentioned the growth rates and you showed some charts at the beginning of the presentation about the accelerating growth rate. Over time, as analytics grows, where do you think that growth rate can get to? Can it go as high as the Risk division? The second question is, Taylor and Francis have a basic journals business, and they are saying that because of open access, their primary research business

can grow at around 4%. So, would you say the same is true for your business, that you can grow the primary research business at around 4%? That's the second question. And then the third question is, in terms of percentage of revenue in the primary research business, what percentage of customers basically are bodies that don't publish anything, so that when you move to a pay for publishing model, that those people won't pay anymore? Thank you.

Kumsal Bayazit:

Okay. Thank you very much, Matthew. So, let me start with the growth rate. My objective is to continue to improve the growth trajectory. And there are three reasons why I'm confident. One is the growth that we're driving with analytics tools, second is the long term strong business fundamentals. We show the growth on R&D expansion, health expansion, the users of our products, researchers, doctors and nurses growing over time, as well as information intensity and data intensity growing over time, which drives demand for analytics. And a third is the shift in business mix, as we shift the business more from print, which has a drag to database and tools and corporate primary research. So, fundamentally, we're driving the long term growth trajectory. It's not about any one year and my bias naturally, as the CEO, is to go as fast as possible, but it takes time as 75% of the business is subscription based in long term contracts.

And it also takes time to get the customer adoption for high value decision tools as well. You have seen where we're heading with database and tools, toward higher value decisions similar to our risk business, and that part of the business will become a larger part of our revenues over time. So, hopefully that helps with that question. In terms of primary research business, the way we think about our primary research is we want to be the high quality provider with the lowest effective unit cost for our customers. So, if volume grows like this, revenue will grow just below that, and then cost usually grows just below that as well. So, we are actually improving the value equation for our primary research customers every single year. That has been our strategy for the last decade, if not longer, and that will continue to be our strategy in the primary research business.

And then in terms of number of institutions we have that do not publish any research at all, that tends to be the exception rather than the rule. Most academic institutions and most research intensive academic institutions will actually publish research.

They may have different rates of publishing, but most of our customers do publish research as well. So, thank you, Matthew.

Matthew Walker: Thank you.

Call Operator: And we have no further questions, so I will hand back to Kumsal for the conclusion.

Kumsal Bayazit: Great. Well, that concludes our session today. Thank you very much for joining. We really appreciate it. Thank you.

ENDS