Abhijit Chakraborty	Testing “efficient supply chain propositions” using topological characterization of the global supply chain network	In this paper, we study the topological properties of the global supply chain network in terms of its degree distribution, clustering coefficient, degree-degree correlation, bow-tie structure, and community structure to test the efficient supply chain propositions proposed by E. J. S. Hearnshaw et al. The global supply chain data in the year 2017 are constructed by collecting various company data from the web site of Standard & Poor’s Capital IQ platform. The in- and out-degree distributions are characterized by a power law of the form of $\gamma_{in} = 2.42$ and $\gamma_{out} = 2.11$. The clustering coefficient decays with an exponent $\beta_k = 0.46$. The nodal degree-degree correlations $\langle knn(k) \rangle$ indicates the absence of assortativity. The bow-tie structure of giant weakly connected component (GWCC) reveals that the OUT component is the largest and consists 41.1% of all firms. The giant strong connected component (GSCC) is comprised of 16.4% of all firms. We observe that upstream or downstream firms are located a few steps away from the GSCC. Furthermore, we uncover the community structures of the network and characterize them according to their location and industry classification. We observe that the largest community consists of the consumer discretionary sector based mainly in the United States (US). These firms belong to the OUT component in the bow-tie structure of the global supply chain network. Finally, we confirm the validity of Hearnshaw et al.’s efficient supply chain propositions, namely Proposition S1 (short path length), Proposition S2 (power-law degree distribution), Proposition S3 (high clustering coefficient), Proposition S4 (“fit-gets-richer” growth mechanism), Proposition S5 (truncation of power-law degree distribution), and Proposition S7 (community structure with overlapping boundaries) regarding the global supply chain network. While the original propositions S1 just mentioned a short path length, we found the short path from the GSCC to IN and OUT by analyzing the bow-tie structure. Therefore, the short path length in the bow-tie structure is a conceptual addition to the original propositions of Hearnshaw.
Abhijit Chakraborty	Is Firm-Level International Trade More Pronounced at the Inter-industry or Intra-industry Level?	Traditionally, imports and exports between developing and developed countries have been considered inter-industry trade in which primary commodities and industrial products are imported and exported, while imports and exports between developed countries are considered intra-industry trade. In recent years, economic integration has led to the development of a form of intra-industry trade that corresponds to the cross-border division of labor between developing countries and developed countries or between developing countries, which has been increasing. We analyze industry sector-specific international trade network and global inter-firm production network to determine whether international trade is more pronounced at inter-industry or intra-industry level. The identified communities in the international trade network reveal a six-backbones structure. We find that each community consists of mainly the same or similar industries. We also find that the first to fifth-largest communities involve developed countries, while the sixth-largest linked community involves only developed countries. This community structure means that international trade is actively transacted among the same or similar industry sectors. Conversely, all communities except for the fifth-largest community in the global production network involve both developed and developing countries. This result is consistent with the results obtained in the sector-specific international trade network. The obtained results support the assertion that firm-level international trade is more pronounced at intra-industry level, which corresponds to the cross-border division of labor between developing countries and developed countries.

Amelie Desvars-Larrive	A structured open dataset of government interventions in response to COVID-19	In response to the COVID-19 pandemic, governments have implemented a wide range of non-pharmaceutical interventions (NPIs). Monitoring and documenting government strategies during the COVID-19 crisis is crucial to understand the progression of the epidemic. Following a content analysis strategy of existing public information sources, we developed a specific hierarchical coding scheme for NPIs. We generated a comprehensive structured dataset of government interventions and their respective timelines of implementation. To improve transparency and motivate collaborative validation process, information sources are shared via an open library. We also provide codes that enable users to visualise the dataset. Standardization and structure of the dataset facilitate inter-country comparison and the assessment of the impacts of different NPI categories on the epidemic parameters, population health indicators, the economy, and human rights, among others. This dataset provides an in-depth insight of the government strategies and can be a valuable tool for developing relevant preparedness plans for pandemic. We intend to further develop and update this dataset until the end of December 2020.
Anna Di Natale	Colexification Networks Encode Affective Meaning	Colexification is a linguistic phenomenon that occurs when multiple concepts are expressed in a language with the same word. Colexification patterns are frequently used to estimate the meaning similarity between words, but the hypothesis that these are related is still missing direct empirical validation at scale. Here, we show for the first time that words linked by colexification patterns capture similar affective meanings. Using pre-existing translation data, we extend colexification databases to cover much longer word lists. We achieve this with an unsupervised method of affective lexicon extension that uses colexification network data to interpolate the affective ratings of words that are not included in the original lexicon. We find positive correlations between network-based estimates and empirical affective ratings, which suggest that colexification networks contain information related to affective meanings. Finally, we compare our network method with state-of-the-art machine learning, trained on a large corpus, and show that our simple linguistics-informed unsupervised algorithm yields comparable performance with high explainability. These results show that it is possible to automatically expand affective norms lexica to cover exhaustive word lists when additional data are available, such as in colexification networks.
Bernhard Haslhofer	GraphSense: A General-Purpose Cryptoasset Analytics Platform	There is currently an increasing demand for cryptoasset analysis tools among cryptoasset service providers, the financial industry in general, as well as across academic fields. At the moment, one can choose between commercial services or low-level open-source tools providing programmatic access. In this paper, we present the design and implementation of another option: the GraphSense Cryptoasset Analytics Platform, which can be used for interactive investigations of monetary flows and, more importantly, for executing advanced analytics tasks using a standard data science tool stack. By providing a growing set of open-source components, GraphSense could ultimately become an instrument for scientific investigations in academia and a possible response to emerging compliance and regulation challenges for businesses and organizations dealing with cryptoassets.
Bernhard Haslhofer	Knowledge Graphs in the Libraries and Digital Humanities Domain	Knowledge graphs represent concepts (e.g., people, places, events) and their semantic relationships. As a data structure, they underpin a digital information system, support users in resource discovery and retrieval, and are useful for navigation and visualization purposes. Within the libaries and humanities domain, knowledge graphs are typically rooted in knowledge organization systems, which have a century-old tradition and have undergone their digital transformation with the advent of the Web and Linked Data. Being exposed to the Web, metadata and concept definitions are now forming an interconnected and decentralized global knowledge network that can be curated and enriched by community-driven editorial processes. In the future, knowledge graphs could be vehicles for formalizing and connecting findings and insights derived from the analysis of possibly large-scale corpora in the libraries and digital humanities domain.

Christian Diem	Quantifying Firm-Level Economic Systemic Risk from Nation-Wide Supply Networks	Crises like COVID-19 or the Japanese earthquake in 2011 exposed the fragility of corporate supply networks. The production of goods and services is a highly interdependent process and can be severely impacted by the default of critical suppliers or customers. While knowing the impact of individual companies on national economies is a prerequisite for efficient risk management, the quantitative assessment of the involved economic systemic risks (ESR) is hitherto practically non-existent, mainly because of a lack of fine-grained data in combination with coherent methods. Based on a unique value added tax dataset we derive the detailed production network of an entire country and present a novel approach for computing the ESR of all individual firms. We demonstrate that a tiny fraction (0.035%) of companies has extraordinarily high systemic risk impacting about 23% of the national economic production should any of them default. Firm size alone cannot explain the ESR of individual companies; their position in the production networks does matter substantially. If companies are ranked according to their economic systemic risk index (ESRI), firms with a rank above a characteristic value have very similar ESRI values, while for the rest the rank distribution of ESRI decays slowly as a power-law; 99.8% of all companies have an impact on less than 1% of the economy. We show that the assessment of ESR is impossible with aggregate data as used in traditional Input-Output Economics. We discuss how simple policies of introducing supply chain redundancies can reduce ESR of some extremely risky companies.
Elma Hot Dervic	Soil data clustering by using K-means and fuzzy K-means algorithm	A problem of soil clustering and spatial representation of the obtained results, based on in-situ measurements of physical and chemical characteristics of soil, is analysed in the paper. K-means and fuzzy K-means algorithms are adapted for the soil data clustering. Database of soil samples sampled in Montenegro is used for comparative analysis of the used algorithm. Classified soil data are presented on static Google map.
Fariba Karimi	Advances in the agent-based modeling of economic and social behavior	In this review we discuss advances in the agent-based modeling of economic and social systems. We show the state of the art of the heuristic design of agents and how behavioral economics and laboratory experiments have improved the modeling of agent behavior. We further discuss how economic networks and social systems can be modeled and we discuss novel methodology and data sources. Lastly, we present an overview of estimation techniques to calibrate and validate agent-based models and show avenues for future research.
Fariba Karimi	Homophily influences ranking of minorities in social networks	Homophily can put minority groups at a disadvantage by restricting their ability to establish links with a majority group or to access novel information. Here, we show how this phenomenon can influence the ranking of minorities in examples of real-world networks with various levels of heterophily and homophily ranging from sexual contacts, dating contacts, scientific collaborations, and scientific citations. We devise a social network model with tunable homophily and group sizes, and demonstrate how the degree ranking of nodes from the minority group in a network is a function of (i) relative group sizes and (ii) the presence or absence of homophilic behaviour. We provide analytical insights on how the ranking of the minority can be improved to ensure the representativeness of the group and correct for potential biases. Our work presents a foundation for assessing the impact of homophilic and heterophilic behaviour on minorities in social networks.
Fariba Karimi	Threshold model of cascades in temporal networks	Threshold models try to explain the consequences of social influence like the spread of fads and opinions. Along with models of epidemics, they constitute a major theoretical framework of social spreading processes. In threshold models on static networks, an individual changes her state if a certain fraction of her neighbors has done the same. When there are strong correlations in the temporal aspects of contact patterns, it is useful to represent the system as a temporal network. In such a system, not only contacts but also the time of the contacts are represented explicitly. There is a consensus that bursty temporal patterns slow down disease spreading. However, as we will see, this is not a universal truth for threshold models. In this work, we propose an extension of Watts' classic threshold model to temporal networks. We do this by assuming that an agent is influenced by contacts which lie a certain time into the past. I.e., the individuals are affected by contacts within a time window. In addition to thresholds as the fraction of contacts, we also investigate the number of contacts within the time window as a basis for influence. To elucidate the model's behavior, we run the model on real and randomized empirical contact datasets.

Frank Neffke	Skill relatedness and firm diversification	Because of the importance of human capital, a firm's choice of diversification targets will depend on whether these targets offer opportunities for leveraging existing human resources. We propose to quantify the similarity of different industries' human capital or skill requirements, that is, the industries' skill relatedness, by using information on cross-industry labor flows. Labor flows among industries can be used to identify skill relatedness, because individuals changing jobs will likely remain in industries that value the skills associated with their previous work. Estimates show that firms are far more likely to diversify into industries that have ties to the firms' core activities in terms of our skill-relatedness measure than into industries without such ties or into industries that are linked by value chain linkages or by classification-based relatedness. Copyright © 2012 John Wiley & Sons, Ltd.
Frank Neffke	The value of complementary co-workers	As individuals specialize in specific knowledge areas, a society's know-how becomes distributed across different workers. To use this distributed know-how, workers must be coordinated into teams that, collectively, can cover a wide range of expertise. This paper studies the interdependencies among co-workers that result from this process in a population-wide dataset covering educational specializations of millions of workers and their co-workers in Sweden over a 10-year period. The analysis shows that the value of what a person knows depends on whom that person works with. Whereas having co-workers with qualifications similar to one's own is costly, having co-workers with complementary qualifications is beneficial. This co-worker complementarity increases over a worker's career and offers a unifying frame-work to explain seemingly disparate observations, answering questions such as "Why do returns to education differ so widely?" "Why do workers earn higher wages in large establishments?" "Why are wages so high in large cities?"
Franz Papst	Embracing Opportunities of Livestock Big Data Integration with Privacy Constraints	Today's herd management undergoes a major transformation triggered by the penetration of cheap sensor solutions into cattle farms, and the promise of predictive analytics to detect animal health issues and product-related problems before they occur. The latter is particularly important to prevent disease spread, ensure animal health, animal welfare and product quality. Sensor businesses entering the market tend to build their solutions as end-to-end pipelines spanning sensors, proprietary algorithms, cloud services, and mobile apps. Since data privacy is an important issue in this industry, as a result, disconnected data silos, heterogeneity of APIs, and lack of common standards limit the value the sensor technologies could provide for herd management. In the last few years, researchers and communities proposed a number of data integration architectures to enable exchange between streams of sensor data. This paper surveys the existing efforts and outlines the opportunities they fail to address by treating sensor data as a black box. We discuss alternative solutions to the problem based on privacy-preserving collaborative learning, and provide a set of scenarios to show their benefits for both farmers and businesses.
Franz Papst	Privacy-preserving machine learning for time series data: PhD forum abstract	Machine learning has a lot of potential when applied to time series sensor data, yet a lot of this potential is currently not utilized, due to privacy concerns of parties in charge of this data. In this work I want to apply privacy-preserving techniques to machine learning for time series data, in order to unleash the dormant potential of this type of data.
Hannah Metzler	Collective Emotions during the COVID-19 Outbreak	The COVID-19 pandemic has exposed the world's population to sudden challenges that elicited strong emotional reactions. Although investigations of responses to tragic one-off events exist, studies on the evolution of collective emotions during a pandemic are missing. We analyzed the digital traces of emotional expressions in tweets during five weeks after the start of outbreaks in 18 countries and six different languages. We observed an early strong upsurge of anxiety-related terms in all countries, which was stronger in countries with stronger increases in cases. Sadness terms rose and anger terms decreased around two weeks later,as social distancing measures were implemented. Positive emotions remained relatively stable. All emotions changed together with an increase in the stringency of measures during certain weeks of the outbreak. Our results show some of the most enduring changes in emotional expression observed in long periods of social media data. Words that frequently occurred in tweets suggesta shift in topics of conversation across all emotions, from political ones in 2019, to pandemic related issues during the outbreak, including everyday life changes, other people, and health. This kind of time-sensitive analyses of large-scale samples of emotional expression have the potential to inform mental health support and risk communication.

Jan Korbel	Modeling of financial processes with a space-time fractional diffusion equation of varying order	In this paper, a new model for financial processes in form of a space-time fractional diffusion equation of varying order is introduced, analyzed,and applied for some financial data. While the orders of the spatial andtemporal derivatives of this equation can vary on different time intervals,their ratio remains constant and thus the global scaling properties of itssolutions are conserved. In this way, the model covers both a possiblecomplex short-term behavior of the financial processes and their long-termdynamics determined by its characteristic time-independent scaling expo-nent. As an application, we consider the option pricing and describe how itcan be modeled by the space-time fractional diffusion equation of varyingorder. In particular, the real option prices of index S&P 500 traded inNovember 2008 are analyzed in the framework of our model and the resultsare compared with the predictions made by other option pricing models.
Jan Korbel	Stochastic thermodynamics and fluctuation theorems for non-linear systems	We extend stochastic thermodynamics by relaxing the two assumptions that the Markovian dynamics must be linear and that the equilibrium distribution must be a Boltzmann distribution. We show that if we require the second law to hold when those assumptions are relaxed, then it cannot be formulated in terms of Shannon entropy. However, thermodynamic consistency is salvaged if we reformulate the second law in terms of generalized entropy; our first result is an equation relating the precise form of the non-linear master equation to the precise associated generalized entropy which results in thermodynamic consistency. We then build on this result to extend the usual trajectory-level definitions of thermodynamic quantities that are appropriate even when the two assumptions are relaxed. We end by using these trajectory-level definitions to derive extended versions of the Crooks fluctuation theorem and Jarzynski equality which apply when the two assumptions are relaxed.
Johannes Sorger	Complexity, transparency and time pressure: practical insights into science communication in times of crisis	A global crisis such as the COVID-19 pandemic that started in early 2020 poses significant challenges for how research is conducted and communicated. We present four case studies from the perspective of an interdisciplinary research institution that switched to “corona-mode” during the first two months of the crisis, focussing all its capacities on COVID-19-related issues, communicating to the public directly and via media, as well as actively advising the national government. The case studies highlight the challenges posed by the increased time pressure, high demand for transparency, and communication of complexity and uncertainty. The article gives insights into how these challenges were addressed in our research institution and how science communication in general can be managed during a crisis.
Johannes Wachs	Complexity science approach to economic crime	János Kertész and Johannes Wachs discuss how complexity science and network science are particularly useful for identifying and describing the hidden traces of economic misbehaviour such as fraud and corruption.
Johannes Wachs	Does crowdfunding really foster innovation? Evidence from the board game industry	Crowdfunding offers inventors and entrepreneurs alternative access to resources with which they can develop and realize their ideas. Besides helping to secure capital, crowdfunding also connects creators with engaged early supporters who provide public feedback. But does this process foster truly innovative outcomes? Does the proliferation of crowdfunding in an industry make it more innovative overall? Prior studies investigating the link between crowdfunding and innovation do not compare traditional and crowdfunded products and so while claims that crowdfunding supports innovation are theoretically sound, they lack empirical backing. We address this gap using a unique dataset of board games, an industry with significant crowdfunding activity in recent years. Each game is described by how it combines fundamental mechanisms such as dice-rolling, negotiation, and resource-management, from which we develop quantitative measures of innovation in game design. Using these measures to compare games, we find that crowdfunded games tend to be more distinctive from previous games than their traditionally published counterparts. They are also significantly more likely to implement novel combinations of mechanisms. Crowdfunded games are not just transient experiments: subsequent games imitate their novel ideas. These results hold in regression models controlling for game and designer-level confounders. Our findings demonstrate that the innovative potential of crowdfunding goes beyond individual products to entire industries, as new ideas spill over to traditionally funded products.

Johannes Wachs	The Geography of Open Source Software: Evidence from GitHub	Open Source Software plays an important role in the digital economy. Yet although software production is amenable to remote collaboration and its end products are easily shared across distances, software development seems to cluster geographically in places such as Silicon Valley, London, or Berlin. And while recent work indicates that positive effects of open source software production accrue locally through knowledge spillovers and information effects, up-to-date data on the geographic distribution of active open source developers remains limited. Here we analyze the geographic distribution of more than half a million active contributors to GitHub located in early 2021 at various spatial scales. Comparing our data with results from before 2010, we find a significant increase in the relative share of developers based in Asia, Latin America and Eastern Europe, suggesting a more even spread of OSS developers globally. Within countries, however, we find significant concentration in regions, exceeding by some margin the concentration of workers in high-tech fields. We relate OSS activity to a number of social and technological indicators at both scales using a multiple regression framework. Despite the potential of OSS as a distributed mode of collaborative work, the data suggest that OSS activity remains highly localized.
Katharina Ledebur	Meteorological factors and non-pharmaceutical interventions explain local differences in the spread of SARS-CoV-2 in Austria	The drivers behind regional differences of SARS-CoV-2 spread on finer spatio-temporal scales are yet to be fully understood. Here we develop a data-driven modelling approach based on an age-structured compartmental model that compares 116 Austrian regions to a suitably chosen control set of regions to explain variations in local transmission rates through a combination of meteorological factors, non-pharmaceutical interventions and mobility. We find that more than 60% of the observed regional variations can be explained by these factors. Decreasing temperature and humidity, increasing cloudiness, precipitation and the absence of mitigation measures for public events are the strongest drivers for increased virus transmission, leading in combination to a doubling of the transmission rates compared to regions with more favourable weather. We conjecture that regions with little mitigation measures for large events that experience shifts toward unfavourable weather conditions are particularly predisposed as nucleation points for the next seasonal SARS-CoV-2 waves.
Marcia Ferreira	Quantifying exaptation in scientific evolution	Rediscovering a new function for something can be just as important as the discovery itself. In 1982, Stephen Jay Gould and Elisabeth Vrba named this phenomenon Exaptation to describe a radical shift in the function of a specific trait during biological evolution. While exaptation is thought to be a fundamental mechanism for generating adaptive innovations, diversity, and sophisticated features, relatively little effort has been made to quantify exap-tation outside the topic of biological evolution. We think that this concept provides a useful framework for characterising the emergence of innovations in science. This article explores the notion that exaptation arises from the usage of scientific ideas in domains other than the area that they were originally applied to. In particular, we adopt a normalised entropy and an inverse participation ratio as observables that reveal and quantify the concept of exapta- tion. We identify distinctive patterns of exaptation and expose specific examples of papers that display those patterns. Our approach represents a first step towards the quantification of exaptation phenomena in the context of scientific evolution.
Marcia Ferreira	Scholars mobility and its impact on the knowledge producers' workforce of European regions	Knowledge production increasingly relies on mobility. However, its role as a mechanism for knowledge recombination and dissemination remains largely unknown. Based on 1,244,080 Web of Science publications from 1,435,729 authors that we used to construct a panel dataset, we study the impact of inter-regional publishing and scientists' mobility in fostering the workforce composition of European countries during 2008-2017. Specifically, we collect information on scientists who have published in one region and then published elsewhere, and explore some determinants of regional and international mobility. Preliminary findings suggest that while talent pools of researchers are increasingly international, their movements seem to be steered by geographical structures. Future research will investigate the impact of mobility on the regional structure of scientific fields by accounting for the appearance and disappearance of research topics over time.

Max Pellert	Validating daily social media macroscopes of emotions	To study emotions at the macroscopic level, affective scientists have made extensive use of sentiment analysis on social media text. However, this approach can suffer from a series of methodological issues with respect to sampling biases and measurement error. To date, it has not been validated if social media sentiment can measure the day to day temporal dynamics of emotions aggregated at the macro level of a whole online community. We ran a large-scale survey at an online newspaper to gather daily self-reports of affective states from its users and compare these with aggregated results of sentiment analysis of user discussions on the same online platform. Additionally, we preregistered a replication of our study using Twitter text as a macroscope of emotions for the same community. For both platforms, we find strong correlations between text analysis results and levels of self-reported emotions, as well as between inter-day changes of both measurements. We further show that a combination of supervised and unsupervised text analysis methods is the most accurate approach to measure emotion aggregates. We illustrate the application of such social media macroscopes when studying the association between the number of new COVID-19 cases and emotions, showing that the strength of associations is comparable when using survey data as when using social media data. Our findings indicate that macro level dynamics of affective states of users of an online platform can be tracked with social media text, complementing surveys when self-reported data is not available or difficult to gather.
Michaela Kaleta	How Specialist Aftercare Impacts Long-Term Readmission Risks in Elderly Patients With Metabolic, Cardiac, and Chronic Obstructive Pulmonary Diseases: Cohort Study Using Administrative Data	The health state of elderly patients is typically characterized by multiple co-occurring diseases requiring the involvement of several types of health care providers. We aimed to quantify the benefit for multimorbid patients from seeking specialist care in terms of long-term readmission risks.From an administrative database, we identified 225,238 elderly patients with 97 different diagnosis (ICD-10 codes) from hospital stays and contact with 13 medical specialties. For each diagnosis associated with the first hospital stay, we used multiple logistic regression analysis to quantify the sex-specific and age-adjusted long-term all-cause readmission risk (hospitalizations occurring between 3 months and 3 years after the first admission) and how specialist contact impacts these risks. Men have a higher readmission risk than women (mean difference over all first diagnoses 1.9%, P<.001), but similar reduction in readmission risk after receiving specialist care. Specialist care can reduce readmission risk by almost 50%. We found the greatest reductions in risk when the first hospital stay was associated with diagnoses corresponding to complex chronic diseases such as acute myocardial infarction (57.6% reduction in readmission risk, SE 7.6% for men [m]; 55.9% reduction, SE 9.8% for women [w]), diabetic and other retinopathies (m: 62.3%, SE 8.0; w: 60.1%, SE 8.4%), chronic obstructive pulmonary disease (m: 63.9%, SE 7.8%; w: 58.1%, SE 7.5%), disorders of lipoprotein metabolism (m: 64.7%, SE 3.7%; w: 63.8%, SE 4.0%), and chronic ischemic heart diseases (m: 63.6%, SE 3.1%; w: 65.4%, SE 3.0%). Specialist care can greatly reduce long-term readmission risk for patients with chronic and multimorbid diseases. Further research is needed to identify the specific reasons for these findings and to understand the detected sex-specific differences.
Nicola Cinardi	A generalised model for asymptotically-scale-free geographical networks	We consider a generalised d-dimensional model for asymptotically-scale-free geographical networks. Central to many networks of this kind, when considering their growth in time, is the attachment rule, i.e. the probability that a new node is attached to one (or more) preexistent nodes. In order to be more realistic, a fitness parameter $\eta_i \in [0,1]$ for each node i of the network is also taken into account to reflect the ability of the nodes to attract new ones. Our d-dimensional model takes into account the geographical distances between nodes, with different probability distribution for η which sensibly modifies the growth dynamics. The preferential attachment rule is assumed to be $\Pi_i \propto k_i \eta_i r^{-\alpha A_{ij}}$ where ki is the connectivity of the ith pre-existing site and αA characterizes the importance of the euclidean distance r for the network growth. For special values of the parameters, this model recovers respectively the Bianconi-Barabási and the Barabási-Albert ones. The present generalised model is asymptotically scale-free in all cases, and its degree distribution is very well fitted with q-exponential distributions, which optimise the nonadditive entropy Sq, given by $p(k) \propto e_{\kappa q}^{-k/\kappa} \equiv 1/[1+(q-1)k/\kappa]^{1/(q-1)}$, with (q,κ) depending uniquely only on the ratio αA/d and the fitness distribution. Hence this model constitutes a realization of asymptotically-scale-free geographical networks within nonextensive statistical mechanics, where k plays the role of energy and κ plays the role of temperature. General scaling laws are also found for q as a function of the parameters of the model.

Niklas Reisz	Monte Carlo modelling of the ImagingRing System–a new method for realistic X-ray distribution	Isocenter positioning uncertainty on the EPID was< 0.1 mm. Beam angle alignment accuracy was 0.3 for both directions. Focal spot alignment accuracy was< 0.5 mm. MLC absolute position accuracy was< 0.5 mm. Gantry isocenter locus had a radius of< 0.4 mm.
Niraj Kushwaha	Machine learning assisted chimera and solitary states in networks	Chimera and Solitary states have captivated scientists and engineers due to their peculiar dynamical states corresponding to co-existence of coherent and incoherent dynamical evolution in coupled units in various natural and artificial systems. It has been further demonstrated that such states can be engineered in systems of coupled oscillators by suitable implementation of communication delays. Here, using supervised machine learning, we predict (a) the precise value of delay which is sufficient for engineering chimera and solitary states for a given set of system's parameters, as well as (b) the intensity of incoherence for such engineered states. Ergo, using few initial data points we generate a machine learning model which can then create a more refined phase plot as well as by including new parameter values. We demonstrate our results for two different examples consisting of single layer and multi layer networks. First, the chimera states (solitary states) are engineered by establishing delays in the neighboring links of a node (the interlayer links) in a 2-D lattice (multiplex network) of oscillators. Then, different machine learning classifiers, K-nearest neighbors (KNN), support vector machine (SVM) and multi-layer perceptron neural network (MLP-NN) are employed by feeding the data obtained from the network models. Once a machine learning model is trained using the limited amount of data, it predicts the precise value of critical delay as well as the intensity of incoherence for a given unknown systems parameters values. Testing accuracy, sensitivity, and specificity analysis reveal that MLP-NN classifier is better suited than Knn or SVM classifier for the predictions of parameters values for engineered chimera and solitary states. The technique provides an easy methodology to predict critical delay values as well as intensity of incoherence for that delay value for designing an experimental setup to create solitary and chimera states.
Pietro Saggese	Identifying the Arbitrageurs on Mt. Gox: First Insights from the Leaked Dataset	We mine the leaked history of trades on Mt. Gox, the dominant Bitcoin exchange from 2011 to early 2014, with the aim of identifying investors who performed two-point arbitrage between Mt. Gox and three other cryptocurrency exchanges (BTC-e, Bitstamp, Bitfinex). Most importantly, the availability of user identifiers per trade allows us to reconstruct the sequence of actions executed by each investor. We match these sequences to ‘ideal’sequences of arbitrage trades, considering the price differences between exchanges and a user-specific estimate of transaction costs. The latter involves a fee model that is inspired by the posted fee schedules and fitted to empirical data. The subset of users whose actual series matches the ideal series best are potential arbitrageurs. Out of about 125,000 users, we identify 2,631 potential arbitrageurs with all three counterpart exchanges, and 14,291 potential arbitrageurs with at least one counterpart exchange. We consider these numbers as upper bounds and argue that post-filtering techniques are required to refine and later validate the results. A preliminary comparison of aggregate statistics between potential arbitrageurs and non-arbitrageurs is given and discussed.
R. M. del Rio-Chanona	Supply and demand shocks in the COVID-19 pandemic: an industry and occupation perspective	We provide quantitative predictions of first-order supply and demand shocks for the US economy associated with the COVID-19 pandemic at the level of individual occupations and industries. To analyse the supply shock, we classify industries as essential or non-essential and construct a Remote Labour Index, which measures the ability of different occupations to work from home. Demand shocks are based on a study of the likely effect of a severe influenza epidemic developed by the US Congressional Budget Office. Compared to the pre-COVID period, these shocks would threaten around 20 per cent of the US economy’s GDP, jeopardize 23 per cent of jobs, and reduce total wage income by 16 per cent. At the industry level, sectors such as transport are likely to be output-constrained by demand shocks, while sectors relating to manufacturing, mining, and services are more likely to be constrained by supply shocks. Entertainment, restaurants, and tourism face large supply and demand shocks. At the occupation level, we show that high-wage occupations are relatively immune from adverse supply- and demand-side shocks, while low-wage occupations are much more vulnerable. We should emphasize that our results are only first-order shocks—we expect them to be substantially amplified by feedback effects in the production network.

Rahim Entezari	Class-dependent Compression of Deep Neural Networks	Today's deep neural networks require substantial computation resources for their training, storage, and inference, which limits their effective use on resource-constrained devices. Many recent research activities explore different options for compressing and optimizing deep models. On the one hand, in many real-world applications, we face the data imbalance challenge, i.e. when the number of labeled instances of one class considerably outweighs the number of labeled instances of the other class. On the other hand, applications may pose a class imbalance problem, i.e. higher number of false positives produced when training a model and optimizing its performance may be tolerable, yet the number of false negatives must stay low. The problem originates from the fact that some classes are more important for the application than others, e.g. detection problems in medical and surveillance domains. Motivated by the success of the lottery ticket hypothesis, in this paper we propose an iterative deep model compression technique, which keeps the number of false negatives of the compressed model close to the one of the original model at the price of increasing the number of false positives if necessary. Our experimental evaluation using two benchmark data sets shows that the resulting compressed sub-networks 1) achieve up to 35% lower number of false negatives than the compressed model without class optimization, 2) provide an overall higher AUC_ROC measure, and 3) use up to 99% fewer parameters compared to the original network.
Rudolf Hanel	Time-Energy Uncertainty Principle for Irreversible Heat Engines	Even though irreversibility is one of the major hallmarks of any real-life process, an actual understanding of irreversible processes remains still mostly semi-empirical. In this paper, we formulate a thermodynamic uncertainty principle for irreversible heat engines operating with an ideal gas as a working medium. In particular, we show that the time needed to run through such an irreversible cycle multiplied by the irreversible work lost in the cycle is bounded from below by an irreducible and process-dependent constant that has the dimension of an action. The constant in question depends on a typical scale of the process and becomes comparable to Planck's constant at the length scale of the order Bohr radius, i.e. the scale that corresponds to the smallest distance on which the ideal gas paradigm realistically applies. This article is part of the theme issue 'Fundamental aspects of nonequilibrium thermodynamics'.
Samuel Martín-Gutiérrez	Recurrent patterns of user behavior in different electoral campaigns: a twitter analysis of the Spanish general elections of 2015 and 2016	We have retrieved and analyzed several millions of Twitter messages corresponding to the Spanish general elections held on the 20th of December 2015 and repeated on the 26th of June 2016. The availability of data from two electoral campaigns that are very close in time allows us to compare collective behaviors of two analogous social systems with a similar context. By computing and analyzing the time series of daily activity, we have found a significant linear correlation between both elections. Additionally, we have revealed that the daily number of tweets, retweets, and mentions follow a power law with respect to the number of unique users that take part in the conversation. Furthermore, we have verified that the topologies of the networks of mentions and retweets do not change from one election to the other, indicating that their underlying dynamics are robust in the face of a change in social context. Hence, in the light of our results, there are several recurrent collective behavioral patterns that exhibit similar and consistent properties in different electoral campaigns.
Sina Sajjadi	Social distancing in pedestrian dynamics and its effect on disease spreading	Nonpharmaceutical measures such as social distancing can play an important role in controlling the spread of an epidemic. In this paper, we use a mathematical model combining human mobility and disease spreading. For the mobility dynamics, we design an agent-based model consisting of pedestrian dynamics with a novel type of force to resemble social distancing in crowded sites. For the spreading dynamics, we consider the compartmental susceptible-exposed-infective (SEI) dynamics plus an indirect transmission with the footprints of the infectious pedestrians being the contagion factor. We show that the increase in the intensity of social distancing has a significant effect on the exposure risk. By classifying the population into social distancing abiders and nonabiders, we conclude that the practice of social distancing, even by a minority of potentially infectious agents, results in a drastic change in the population exposure risk, but it reduces the effectiveness of the protocols when practiced by the rest of the population. Furthermore, we observe that for contagions for which the indirect transmission is more significant, the effectiveness of social distancing would be reduced. This study can help to provide a quantitative guideline for policy-making on exposure risk reduction.

Stefan Thurner	The role of mainstreamness and interdisciplinarity for the relevance of scientific papers	Is it possible to tell how interdisciplinary and out-of-the-box scientific papers are, or which papers are mainstream? Here we use the bibliographic coupling network, derived from all physics papers that were published in the Physical Review journals in the past century, to try to identify them as mainstream, out-of-the-box, or interdisciplinary. We show that the network clusters into scientific fields. The position of individual papers with respect to these clusters allows us to estimate their degree of mainstreamness or interdisciplinarity. We show that over the past decades the fraction of mainstream papers increases, the fraction of out-of-the-box decreases, and the fraction of interdisciplinary papers remains constant. Studying the rewards of papers, we find that in terms of absolute citations, both, mainstream and interdisciplinary papers are rewarded. In the long run, mainstream papers perform less than interdisciplinary ones in terms of citation rates. We conclude that to avoid a unilateral trend towards mainstreamness a new incentive scheme is necessary.
Stefan Thurner	Leverage causes fat tails and clustered volatility	We build a simple model of leveraged asset purchases with margin calls. Investment funds use what is perhaps the most basic financial strategy, called ‘value investing’, i.e. systematically attempting to buy underpriced assets. When funds do not borrow, the price fluctuations of the asset are approximately normally distributed and uncorrelated across time. This changes when the funds are allowed to leverage, i.e. borrow from a bank, which allows them to purchase more assets than their wealth would otherwise permit. During good times, funds that use more leverage have higher profits, increasing their wealth and making them dominant in the market. However, if a downward price fluctuation occurs while one or more funds is fully leveraged, the resulting margin call causes them to sell into an already falling market, amplifying the downward price movement. If the funds hold large positions in the asset, this can cause substantial losses. This in turn leads to clustered volatility: before a crash, when the value funds are dominant, they damp volatility, and after the crash, when they suffer severe losses, volatility is high. This leads to power-law tails, which are both due to the leverage-induced crashes and due to the clustered volatility induced by the wealth dynamics. This is in contrast to previous explanations of fat tails and clustered volatility, which depended on ‘irrational behavior’, such as trend following. A standard (supposedly more sophisticated) risk control policy in which individual banks base leverage limits on volatility causes leverage to rise during periods of low volatility, and to contract more quickly when volatility becomes high, making these extreme fluctuations even worse.
Taun Pham	Balance and fragmentation in societies with homophily and social balance	Recent attempts to understand the origin of social fragmentation on the basis of spin models include terms accounting for two social phenomena: homophily—the tendency for people with similar opinions to establish positive relations—and social balance—the tendency for people to establish balanced triadic relations. Spins represent attribute vectors that encode G different opinions of individuals; social interactions between individuals can be positive or negative. Here we present a co-evolutionary Hamiltonian framework that minimizes individuals’ social stress in social networks that have finite connectivity and people with a small number of attributes. We show that such systems always reach stationary, balanced, and fragmented states, if–in addition to homophily–individuals take into account a significant fraction, q, of their triadic relations. Above a critical value, qc, balanced and fragmented states exist for any number of opinions.
Vito DP Servedio	A new and stable estimation method of country economic fitness and product complexity	We present a new metric estimating fitness of countries and complexity of products by exploiting a non-linear non-homogeneous map applied to the publicly available information on the goods exported by a country. The non homogeneous terms guarantee both convergence and stability. After a suitable rescaling of the relevant quantities, the non homogeneous terms are eventually set to zero so that this new metric is parameter free. This new map almost reproduces the results of the original homogeneous metrics already defined in literature and allows for an approximate analytic solution in case of actual binarized matrices based on the Revealed Comparative Advantage (RCA) indicator. This solution is connected with a new quantity describing the neighborhood of nodes in bipartite graphs, representing in this work the relations between countries and exported products. Moreover, we define the new indicator of country net-efficiency quantifying how a country efficiently invests in capabilities able to generate innovative complex high quality products. Eventually, we demonstrate analytically the local convergence of the algorithm involved.

Vito DP Servedio	The dynamics of correlated novelties	Novelties are a familiar part of daily life. They are also fundamental to the evolution of biological systems, human society and technology. By opening new possibilities, one novelty can pave the way for others in a process that Kauffman has called “expanding the adjacent possible”. The dynamics of correlated novelties, however, have yet to be quantified empirically or modeled mathematically. Here we propose a simple mathematical model that mimics the process of exploring a physical, biological, or conceptual space that enlarges whenever a novelty occurs. The model, a generalization of Polya's urn, predicts statistical laws for the rate at which novelties happen (Heaps' law) and for the probability distribution on the space explored (Zipf's law), as well as signatures of the process by which one novelty sets the stage for another. We test these predictions on four data sets of human activity: the edit events of Wikipedia pages, the emergence of tags in annotation systems, the sequence of words in texts and listening to new songs in online music catalogues. By quantifying the dynamics of correlated novelties, our results provide a starting point for a deeper understanding of the adjacent possible and its role in biological, cultural and technological evolution.
William Schueller	Active learning strategies and active control of complexity growth in naming games	Naming Games are models of the dynamic formation of lexical conventions in populations of agents. In this work we introduce new Naming Game strategies, using developmental and active learning mechanisms to control the growth of complexity. An information theoretical measure to compare those strategies is introduced, and used to study their impact on the dynamics of the Naming Game.