Extended Abstract for OEE3 Workshop (ALife Conference 2018 Tokyo)

Detecting the on-going emergence of technological innovations

Mark A. Bedau[†]*, Nicholas Gigliotti[†], Tobias Janssen[†], Alec Kosik[†], Ananthan Nambiar[†], Norman Packard[‡]

[†]Reed College, Portland

[‡]ProtoLife, Inc., San Francisco

*contact author: mab@reed.edu

Abstract

We detect a form of open-ended evolution (OEE) in empirical data about human technological innovations. This evolution occurs in a non-biological, cultural population that exists in the real world and is evolving in a way that seems as open-ended as biological evolution. Using patented inventions as a proxy for technological innovations, we mine public patent records for evidence of one specific form of OEE—the on-going emergence of technological innovations—and we compare two ways to detect it. One way is to detect the first instances of pre-defined patent pigeon holes, such as the technology classes listed in the United States Patent Classification (USPC). The second way is to embed patents in a high-dimensional semantic space and detect the emergence of new clusters. Both methods reveal the on-going generation of new kinds of technologies when applied to hundreds of years of patent records, but only clusters reveal innovations that are unanticipated. Our methodology easily generalizes to detecting unanticipated innovations in other evolving populations that leave rich digital traces.

Keywords: patent, invention, innovation, technology, classification, taxonomy, open-ended evolution

Contents

1 Detecting innovations with pigeon holes and clusters

We develop and apply methods for detecting a kind of open-ended evolution (OEE) in empirical data generated by evolving systems. The methods apply to both models and real systems in the natural world. Discussions of OEE typically focus on biological examples, but central here is a non-biological population in the real world: human technological innovations. Human technology and living organisms differ in many respects, but both exhibit open-ended evolution and the study of each will illuminate the other.

We study technological innovations by means of a convenient proxy: patented inventions. Patent records contain detailed, accurate, digital descriptions of each invention. In the aggregate this big data stream can provide an illuminating window on the evolution of technology. Contemporary methods in statistics and machine learning can reveal a wealth of patterns in the evolution of technology, including our primary focus here: the on-going generation of new innovations. Our methods easily generalize to the detection of innovations in other populations with sufficient digital traces.

Open-ended evolution is widely recognized to be enigmatic (Taylor et al. (2016)) and one reason is the *emergence problem*: the difficulty of detecting entities that are so novel that we have no distinctive descriptions for them (Bedau et al. (1998)). This problem is especially acute for those aiming to detect the emergence of new technological innovations armed only with a pre-defined classification created by human experts. The problem is not just the biases, preconceptions, and other epistemic short-comings of any human expert. A more specific challenge is to create a classification that detects innovations when they were *unanticipated* when they occurred. Recent examples include nanotechnology and genetic medicine. Lacking an accurate classification blinds us to some innovations and increases the number of undetected genuine innovations (false negatives).

Our goal here is one of the forms of OEE listed by Taylor et al. (2016)—the on-going generation of new kinds of entities—and we aim to compare two methods for detecting it. One method classifies entities by distributing them into a fixed and finite list of pre-defined *pigeon holes* and detecting when pigeon holes are first exemplified. The list of pigeon holes is revised from from time to time, in order to classify especially novel innovations properly. The second method classifies entities into connected groups of *clusters* in an abstract technology feature space learned from a huge corpus of many millions of documents. Here we compare these methods by examining hundreds of years of US patent records, with special attention to the period following 1976.

Lists of pigeon holes are constructed by human experts after sifting through a wealth of historical information, and the resulting classifications are often useful for describing historical patterns of innovation. A pigeon hole must be defined before anyone can classify things with it, even if those things pre-date the formulation of those definitions. We will examine whether this historical orientation of pigeon holes blunts their usefulness for detecting unanticipated future innovations.

In both cases we can formally construe a (non-hierarchical) classification as a set of classes of patents, where each class may be labeled with an integer, so the classification is a map from the set of patents to integers $C : P \to I$. We will assume that there are a finite number of different classes of patents. Class i is the set of all patents mapping to i, or $C^{-1}(i)$. For classifications consisting of pigeon holes, $C(i)$ is just an integer indexing the pigeon hole on a pre-defined list. When we classify patents by mapping them into a group of nearby clusters (described below, $C(i)$ is the i^{th} group of cluster centroids. Each group of cluster centroids defines to a locally-connected sub-region in an abstract technology feature space described below.

Evolutionary activity statistics have been used to measure and compare the rate of adaptive innovation in various computer models (Bedau and Packard (1991); Bedau et al. (1998)) as well as in biological populations (Bedau et al. (1997)) and cultural populations (Buchanan et al. (2011)), and we detect innovations with these statistics for both pigeon holes and clusters. The *evolutionary activity* of a pigeon hole or cluster i at time t is defined as simply the cumulative sum of the number of instances of i from its first exemplification to time t. Our use of activity statistics is supported by the hypothesis that new adaptive innovations can be identified by their unusually high evolutionary activity. When evolutionary activity is plotted as a function of time, new adaptive innovations produce steeply rising "waves" and the rate of adaptive innovation can be measured as the rate of production of new activity waves.

2 Innovations detected with technology pigeon holes

The United States Patent and Trademark Office (USPTO) classifies each patent by putting it into one (or more) classes chosen from the United States Patent Classification (USPC). Those classes are further sub-divided into codes and sub-codes, and the classes are further collected into sub-categories and categories, but we can ignore those details here. The USPTO revises the USPC from time to time, and when the precision matters we specifically refer to the USPC classification in effect in 2018, denoted USPC_{2018}. Of course, the USPTO will likely revise the USPC_{2018} again sometime in the future. Looked at historically, the USPC is a time-indexed sequence of classifications, produced and incrementlly revised now and then by human experts.

Figure 1 depicts the evolutionary activity of each class in USPC_{2018}. The bottom figure blows up the activity scale on the y-axis by over two orders of magnitude, in order to highlight the new activity waves caused when classes of technology are first exemplified. The blow-up clearly shows new waves of evolutionary activity continually sweeping up through the figure—the signature of the on-going generation of new significant technological innovations. The density of new waves at a particular time is a measure of the rate of innovation at that time (Bedau et al. (1997)).

Figure 1 shows that many classes were first exemplified early on in the 1850s, but the density of newly exemplified clusters generally lessens over the first hundred years. After 1945 the rate picks up again and large number of classes are first-exemplified at the creation of the modern-day USPTO in 1976. But then rate quickly drops to nothing and remains at zero through the final twenty years. (As an aside, note that many of the flatlined classes in Figure 1 were first exemplified very long ago.)

The continual production of new activity waves like those in Figure 1 provides one window on the contingent and open-ended process by which the technology pigeon holes in USPC_{2018} were exemplified. But pigeon holes and any fixed and finite classification share an obvious limitation: increasing blindness to new innovations as more and more pigeon holes are exemplified. After the last pigeon hole has been exemplified, no further innovations can be detected—not without defining some new pigeon holes (and possibly abandoning some existing ones).

For this very reason the USPTO sometimes revises the USPC. For example, 40% of all US patents issued in 1976 have by now been reclassified (Lafond and Kim (2017)). With time any fixed and finite classification becomes increasing blind to innovations; this probably explains the general decling in the rate of newly exemplified classes in Figure 1 and especially the absence of new activity waves in the last twenty years. If any genuine innovations arose after 1995, they failed to

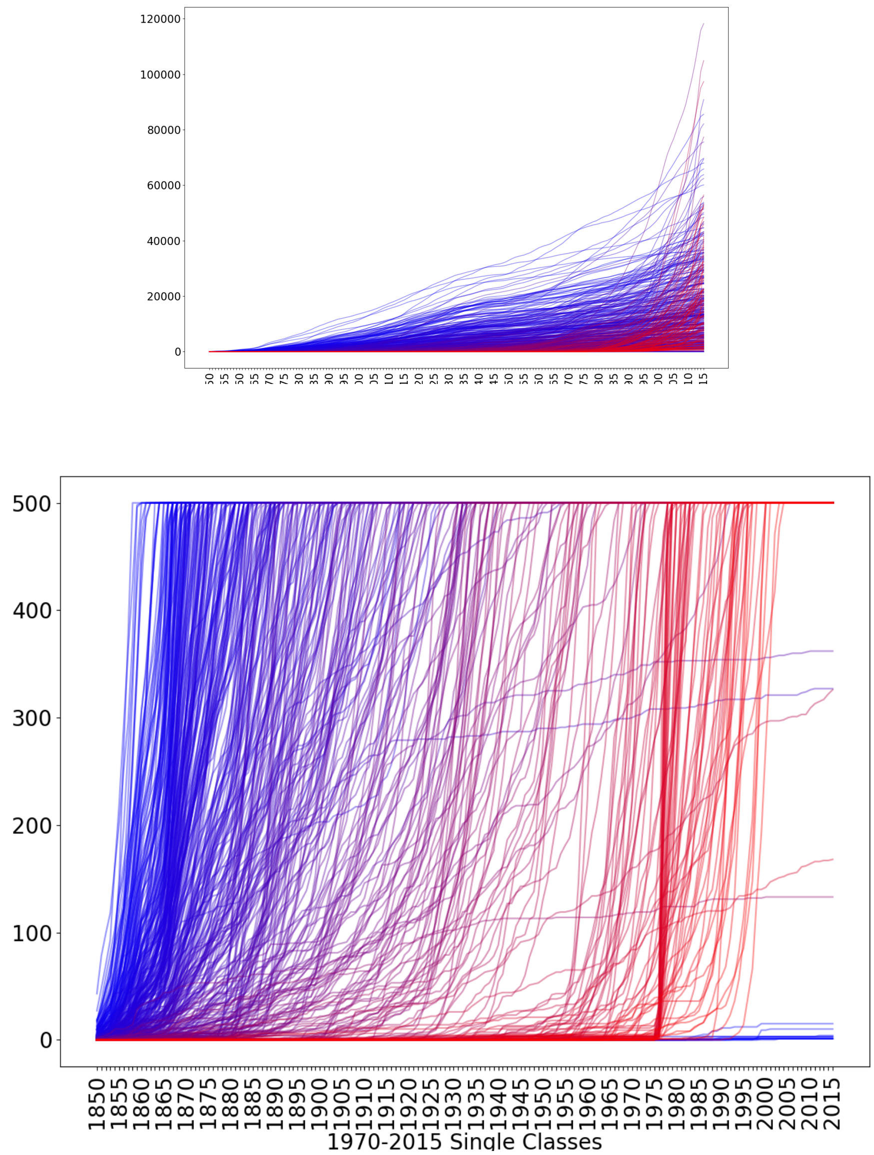

Figure 1: Time series of the evolutionary activity of each class in the USPC$_{2018}$, computed from 1845 to 2015. Below: Blow-up of the bottom 0.5% of the evolutionary activity scale. The frame containing a class's first exemplification is indicated by its color, from oldest (blue) to newest (red).

first exemplify any USPC_{2018} class, and that makes them exactly the sort of false negatives created by the emergence problem.

3 Innovations detected with technology clusters

A different, implicit description of technological innovations emerges empirically from word co-occurrence statistics gleaned by textual analysis of millions of documents describing all of the patents issued by the USPTO since 1850. Packard et al. (2018) recently demonstrated a new classification of technology that evolves over time. The method uses publicly available topic modeling software, `doc2vec` (Le and Mikolov (2014), Rehurek and Sojka (2010)), to build a 300-dimensional semantic vector space. Patent descriptions are embedded in this space with an algorithm that tries to make the *proximity* of two patent documents in the space proportional to the *similarity* of the two inventions described, and a metric on the technology space provides a precise, quantitative measure of the similarity of any two inventions. This embedding space functions as a technology feature space, and different regions in it correspond to different kinds of technologies. (See the references above for more details.)

We collect US patent records from 1976-2014 into temporal "frames" containing 50k successively issued patents, each precisely located in technology space. Then we cluster the patents in each frame using the k-means algorithm, yielding 100 successive frames each containing 25 clusters of patents. The clusters in a frame are stamped with the frame's time index, and different frames can contain clusters in rather different locations, so we can observe where clusters move between frames. The result is a precise description of how clusters in technology space *change and evolve* over time.

To produce an overall classification from these 100 frames with 25 clusters apiece, we collect nearby clusters into *groups*, and we classify a patent into the group that contains the cluster that contains the patent. These groups are defined by a set of centroids that are near one another in technology space; specifically, every centroid in a group is near to at least one other centroid in the group. Consider the network in which the nodes are the 2500 centroids in 100 frames each with 25 centroids, and in which two nodes are connected if and only if they are located in technology space within some pre-defined distance threshold (d_{th}).

Our cluster groups are simply the connected components in this network. Each cluster group picks out a distinctive local region of technology space, defined by clusters in the group. By definition every centroid in a group is *spatially* near some other centroid in the group, but two centroids in the same group could be quite distant spatially provided they are connected by a finite sequence of pairs of nearby centroids. Furthermore, centroids in a group can be *temporally* distant from *every* other centroid in the group; i.e., connections among centroids are atemporal (i.e., temporally "global"). Since groups of centroids are parameterized by d_{th}, changing this parameter can change the groups.

Every patent issued between 1976 and 2014 falls into exactly one of the 25 clusters in one of the 100 data frames. We classify a patent with (index of) the group of clusters that contains the cluster that contains the patent. Since group of clusters grow and change over time, so does the classification they define. Different groups can be exemplified at different times, and over time groups can come into existence, split, merge, and go out of existence. They can also persist over time while moving through technology space. This classification evolves with each crop of new innovations, and on-going analysis of streams of empirical data automatically documents the changing classification.

Relatively few biases, preconceptions, and other human epistemic limitations constrain the construction of statistically-defined technology clusters. The emergence of technology clusters presupposes only a suitable technology space, constructed from word patterns in millions of patent records. Those patent records were authored by millions of different human beings, and those authors no doubt each have individual epistemic limitations. But the construction of technology space averages out those individual quirks. On the other hand, our technology clusters could still reflect any epistemic limitations that millions of patent authors might share.

We locate each cluster in technology space by computing its centroid, and we measure the distance between clusters as the distance between their centroids. It is difficult to visualize high-dimensional clusters, so Figure 2 is a two-dimensional projection of our 2500 cluster centroids, produced with the t-SNE algorithm (Maaten and Hinton (2008)); t-SNE does an especially good job of reflecting local structure in very high dimensional spaces. Each cluster centroid in each frame is shown as a dot, and the time of each frame is encoded in its dot's gray scale, from white (earlier) to black (later). The successive locations of certain centroids describe the movement of the clusters in technology space.

The t-SNE projection in Figure 2 shows that the centroids are located in technology space in readily identifiable groups. Some groups persist through every frame and others first emerge in some later frame. Also, some groups split apart over time and others move closer and merge. These categories of trajectories in technology space indicate the on-going emergence of new clusters in technology space—a form of OEE.

Figure 3 shows all the centroids in the thirty largest groups of clusters, with the columns corresponding to frames of data and the rows corresponding to groups. The groups are ranked down the page from largest to smallest. The largest groups at the top persist through all or most of the frames, and many have more that one

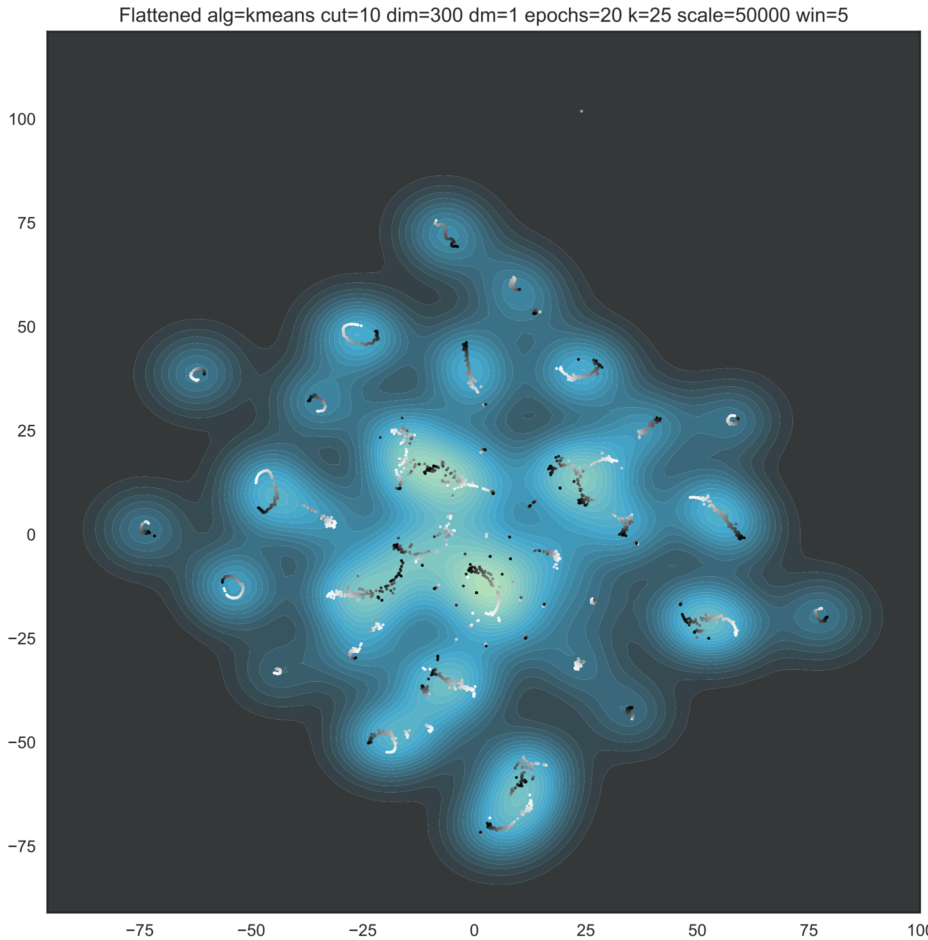

Figure 2: A two-dimensional t-SNE projection of the centroids of clusters of patents issued during the years 1976-2014. Centroids are nearby in a t-SNE projection if and only if they are nearby in the 300-dimensional technology space. Faint contour lines and the landscape color gradient indicate the kernel density estimation for the centroids. One hundred frames of 25 centroids are overlaid. Centroids in each frame have the same color, and gray scale represents time; earlier (later) centroids are lighter (darker).

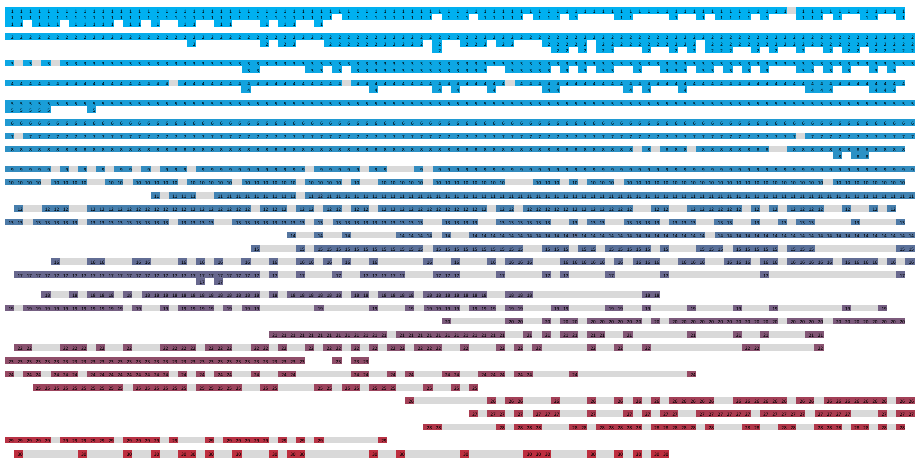

Figure 3: The thirty largest group of clusters, ranked by size and numbered from top to bottom, and displayed in one hundred frames (columns, numbered left to right). Centroids in the same group of clusters have the same color, but connections among centroids within a group are not shown.

centroid in some frames. Groups 2, 5, and 6 have at least one centroids in every single frame, but most groups have temporal gaps in the group's exemplification, and gap frequency and size rises, as expected, with falling group size. The group exemplification patterns in Figure 3 contain some clear examples of the emergence of new innovations (groups 11, 14, 15, 20, 21, 26-28) and the death of old innovations (groups 23, 25, 29, 30). This pattern of exemplification gaps, and births and deaths appears to be robust; even if meta-parameters like size of time frames or number of clusters are changed, the generic pattern remains the same.

The visible groups of centroids in the t-SNE projection (Figure 2) roughly correspond to the groups. Figure 4 shows the thirty largest centroids in Figure 2 and colors them by the identity of the groups in Figure 3 that contain them. The correspondence between cluster groups and the visible groups in the t-SNE confirms that the clusters of patents in technology space arise in distinct local regions.

Figure 5 (top) shows the evolutionary activity of each group of technology clusters that arose during after 1976 and eventually contains at least 4 centroids. The figure (middle) clearly shows the on-going generation of new innovations, as pioneering clusters of patents arise in unexplored regions of technology space and new activity waves start accumulating exemplifications. The distance distributions

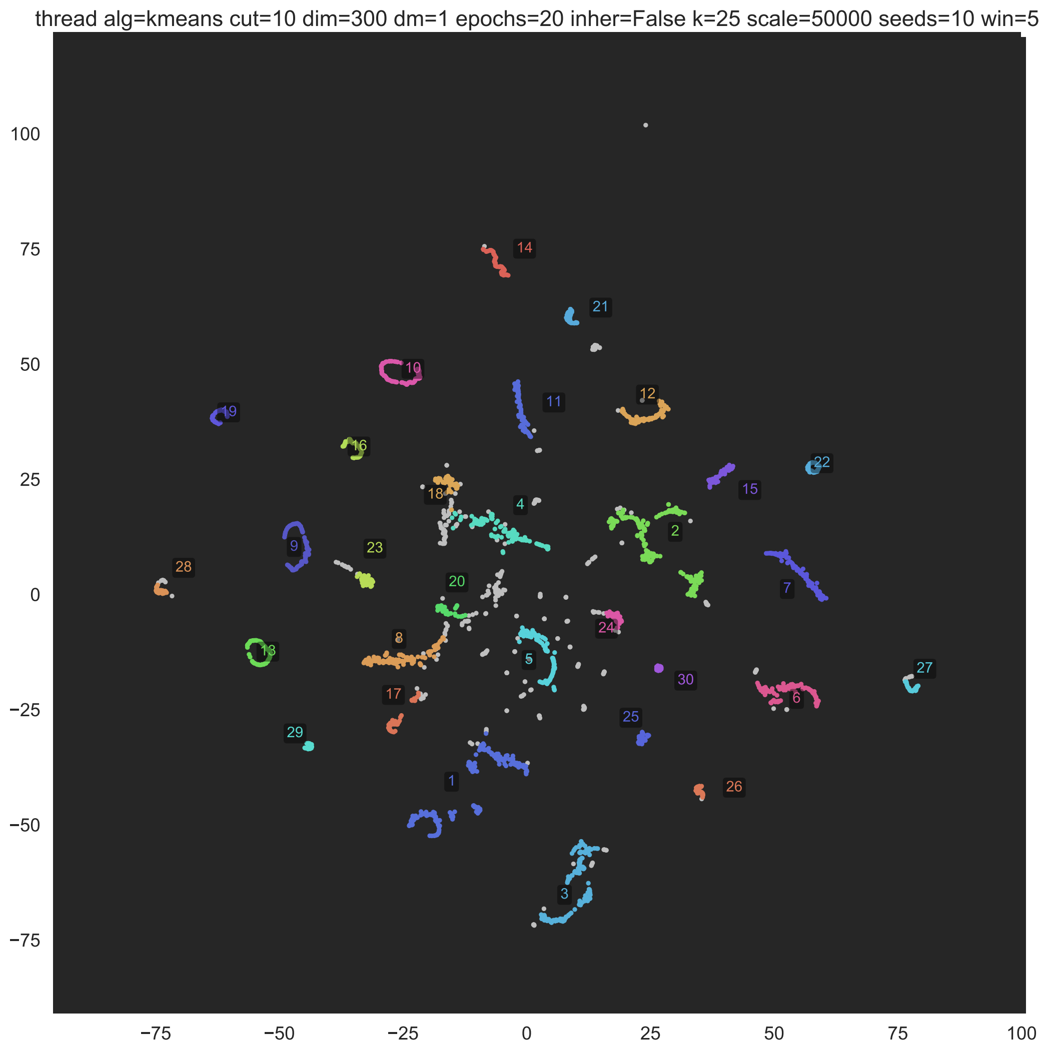

Figure 4: The centroids in the t-SNE projection in Figure 2 are colored and numbered with the thirty largest groups of clusters.

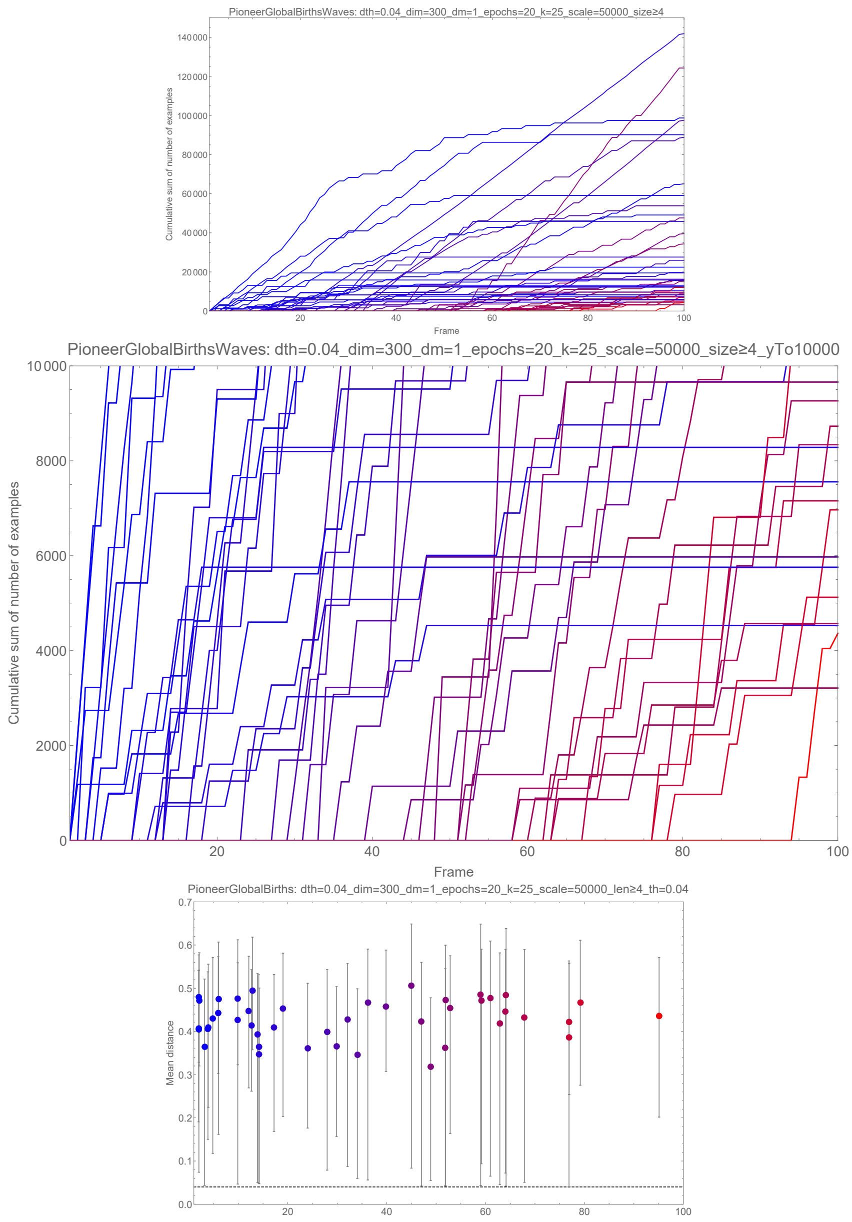

Figure 5: Top: Evolutionary activity of groups of centroids that are first exemplified after the first frame and eventually contain at least 4 centroids. Middle: Blow-up of the bottom 10% of the y-axis above to see the emergence of activity waves that indicate new innovative groups of centroids. Bottom: Temporal sequence of distributions of distances (min, mean, and max) between the earliest centroid in a group of centroids and all other centroids in the same or earlier frames. The black dashed line shows $d_{th} = 0.04$, the distance threshold used to connect centroids into groups of centroids.

in Figure 5 (bottom) show that many of those innovations arise in remote regions of technology space. By construction, any centroid in a group of centroids is farther away than d_{th} from all centroids outside the group. The distance distributions in Figure 5 (bottom) show that about a quarter of the innovations arise in remote regions of technology space, farther away than $4 \times d_{th} = 0.16$ from any centroid in any other group.

Together, Figures 2, 3, and 5 present clear and detailed empirical evidence for the on-going generation of new and persisting clusters of technology—a clear example of technology's open-ended evolution. Figure 5 (middle and bottom) shows dozens of significant innovations that arise after frame 30, which contains clusters of patents issued in part of 1995. But Figure 1 shows that no USPC_{2018} classes were first exemplified after 1995; that is, the 450 pre-defined classes in the USPC_{2018} detect none of the innovations in Figure 5 (middle and bottom) that arise after frame 30. The innovations are all false negatives for the USPC_{2018}.

4 Detecting unanticipated innovations

We have shown how to classify an evolving stream of patent data by first embedding successive frames of patents in technology space, then clustering each frame of patents, and finally grouping centroids of nearby clusters. This is a practical and feasible solution to the problem of detecting and classifying new and emerging technological innovations. The resulting classification is pragmatic and it will continually and automatically adapt to unpredictable changes in the incoming stream of data. Significantly new kinds of inventions will produce new groups of patents that can be revealed by clustering and other adaptive learning algorithms. Our method of detecting new technological innovations is a simple illustration of a very general way to solve the "emergence" problem for open-ended evolution.

Our methods solve analogous emergence problems for many other kinds of innovations such as those happening in biology, chemistry, culture, and beyond. Up until now it has been difficult for observers of dynamically evolving systems to detect novel innovations before the innovation's distinctive characteristics have been been discovered and mapped. Our methods now enable emergence problems in all of these areas to be solved.

The strategy connecting these solutions is effective because it automatically adapts whenever new clusters emerge from the incoming stream of data. Piping a changing real-time data stream through a learning algorithm enables the automatic detection and characterization of unanticipated innovations. These achievements open the door to new methods for forecasting innovations in a wide variety of fields (Packard et al. (2018)).

References

Bedau, M. A. and Packard, N. H. (1991). Measurement of evolutionary activity, teleology, and life. In Langton, C., Taylor, C., Farmer, D., and Rasmussen, S., editors, *Artificial Life II*, pages 431–461. Addison-Wesley.

Bedau, M. A., Snyder, E., Brown, C. T., and Packard, N. H. (1997). A comparison of evolutionary activity in artificial evolving systems and the biosphere. In Husbands, P. and Harvey, I., editors, *Proceedings of the Fourth European Conference on Artificial Life, ECAL97*, pages 125–134. MIT Press.

Bedau, M. A., Snyder, E., and Packard, N. H. (1998). A classification of long-term evolutionary dynamics. In Adami, C., Belew, H., Kitano, H., and Taylor, C., editors, *Artificial Life VI*, pages 228–237. MIT Press.

Buchanan, A., Packard, N. H., and Bedau, M. A. (2011). Measuring the drivers of technological innovation in the patent record. *Artificial Life*, 17(2):109–122.

Lafond, F. and Kim, D. (2017). Long-run dynamics of the U.S. patent classification system. *arXiv:1703.02104v1*.

Le, Q. and Mikolov, T. (2014). Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning (ICML-14)*, pages 1188–1196.

Maaten, L. v. d. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, 9(Nov):2579–2605.

Packard, N. H., Glgliotti, N., Nambiar, A., Janssen, T., and Bedau, M. A. (2018). An evolving classification for forecasting technology. In preparation.

Rehurek, R. and Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*. Citeseer.

Taylor, T., Bedau, M., Channon, A., Ackley, D., Banzhaf, W., Beslon, G., Dolson, E., Froese, T., Hickinbotham, S., Ikegami, T., McMullin, B., Packard, N., Rasmussen, S., Virgo, N., Agmon, E., Clark, E., McGregor, S., Ofria, C., Ropella, G., Spector, L., Stanley, K. O., Stanton, A., Timperley, C., Vostinar, A., and Wiser, M. (2016). Open-ended evolution: Perspectives from the oee workshop in york. *Artificial Life*, 22:408–423.