CZ4125 Developing Data Products

Individual Project Part II:

SCSE Researcher and Faculty Dashboarding

By: Tan Ching Fhen

Mat. ID: U1920787D

Individual SCSE Researcher's Profile Dashboard

The dashboard that I created is displayed in Figures 1 to 3. The left-hand panel shows some basic information about the SCSE Researcher, while the right-hand panel displays a little more detailed information about the quality and quantity of a researcher's research. Both left and right-hand panels can be scrolled down as can be seen in the transition from Figures 1 to 3. Note that every SCSE researcher has his/her own dashboard, which is created by inputting the name in the source code. I used a Python Library, *Panel*, to create these dashboards.

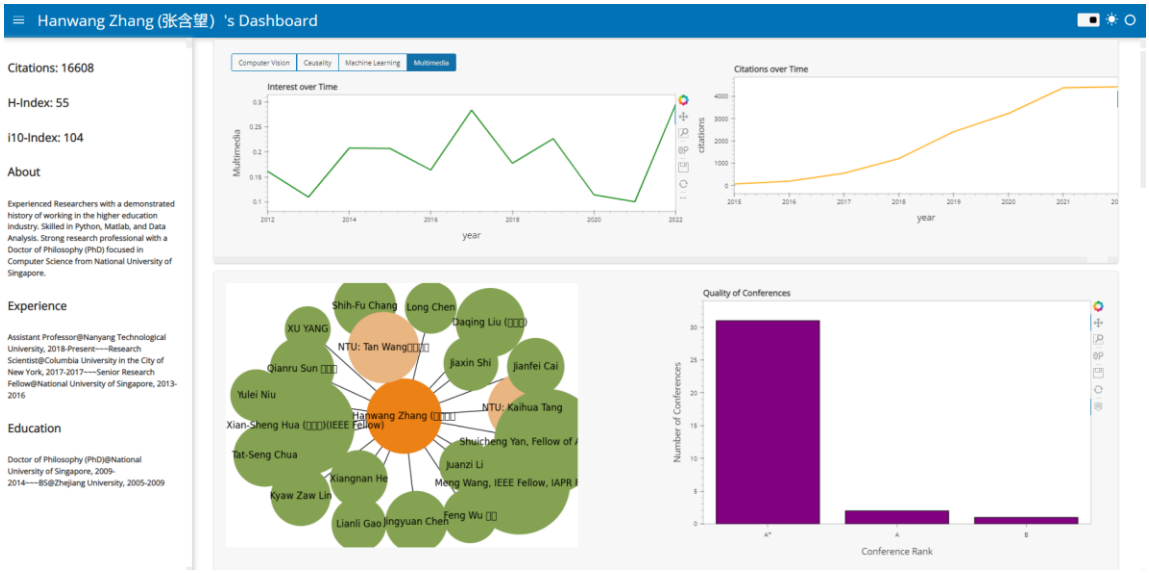


Figure 1 - Hanwang Zhang's Dashboard Part 1

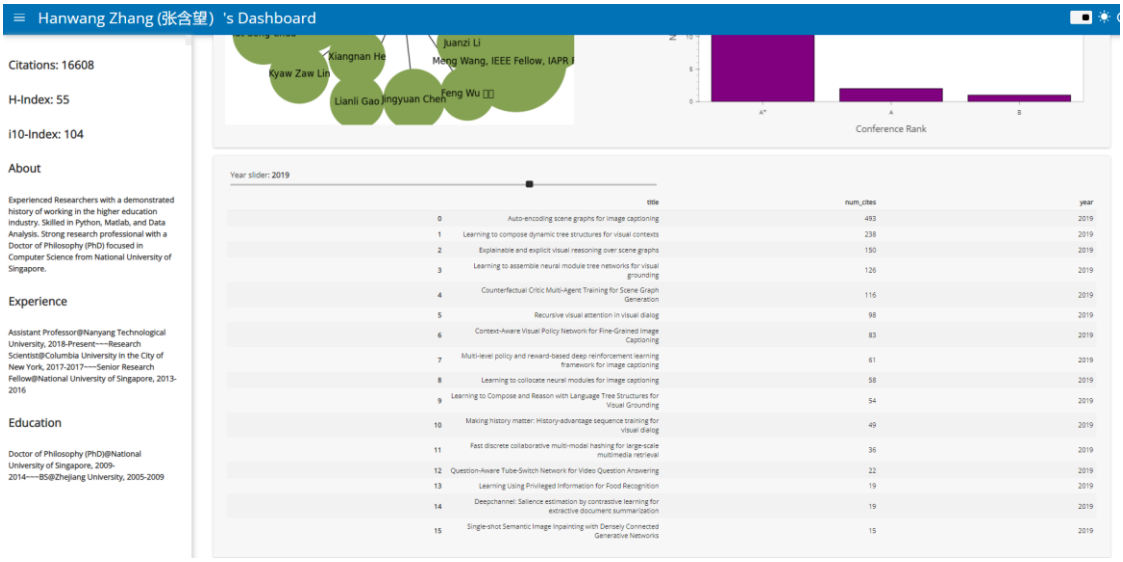


Figure 2 - Hanwang Zhang's Dashboard Part 2

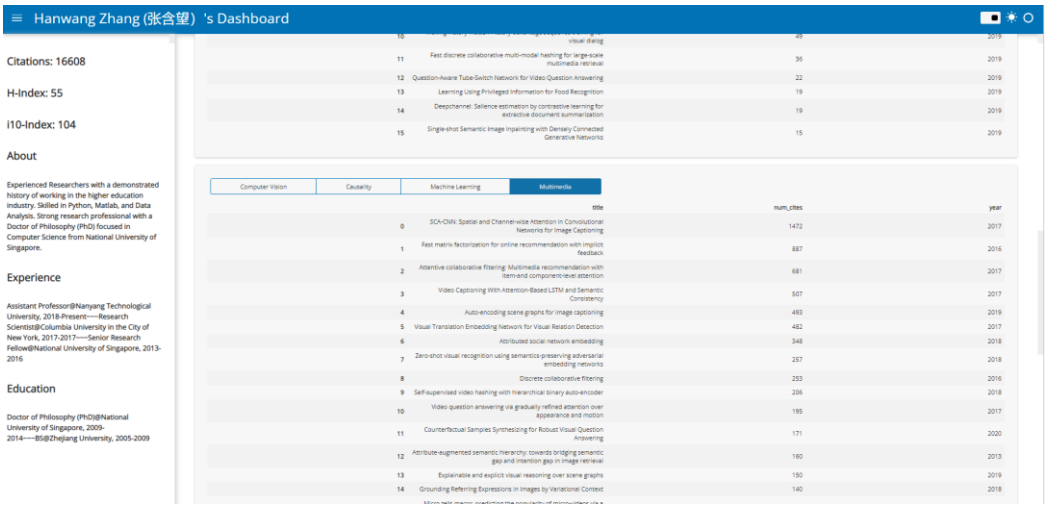


Figure 3 - Hanwang Zhang's Dashboard Part 3

Into more detail, the left-hand panel displays the most useful information first, namely: Number of citations, H-Index (quality of publications), I10-Index (number of publications with more than 10 citations). Following which, I display the summary description of the researcher, experience history and education history (Figures 4 and 5). All these information was either requested from LinkedIn API or Google Scholar Authors API.

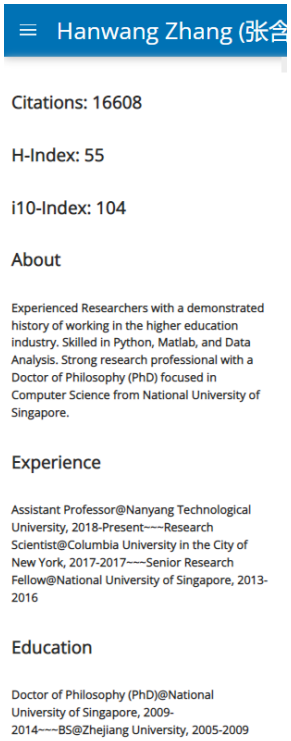


Figure 4 - Hanwang Zhang's basic information on left panel

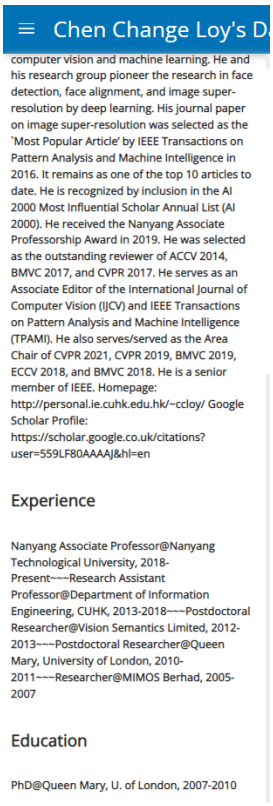


Figure 5 - Chen Change Loy's Basic information on left panel

Moving on to the right-hand panel, the first plot I display is the interest over time. From Figure 6, we can observe that Hanwang Zhang has an increasing interest in multimedia from 2012 to 2017. Following that, a steady decline, and he has the highest interest in multimedia in 2022. On the other hand, his interest in computer vision grew from 2012 to 2017, followed by a steady decline ever since (Figure 7). These interest tags (computer vision, causality, machine learning and multimedia) were from the researcher's Google Scholar Profile. For different researchers, the interest tags are different, depending on what they had put up on the google scholar profiles. For instance, for Cheng Change Loy, he has a different set of interests: computer vision, image processing and machine learning (Figure 8).

Figure 6 - Hanwang Zhang's Interest over time plot - multimedia

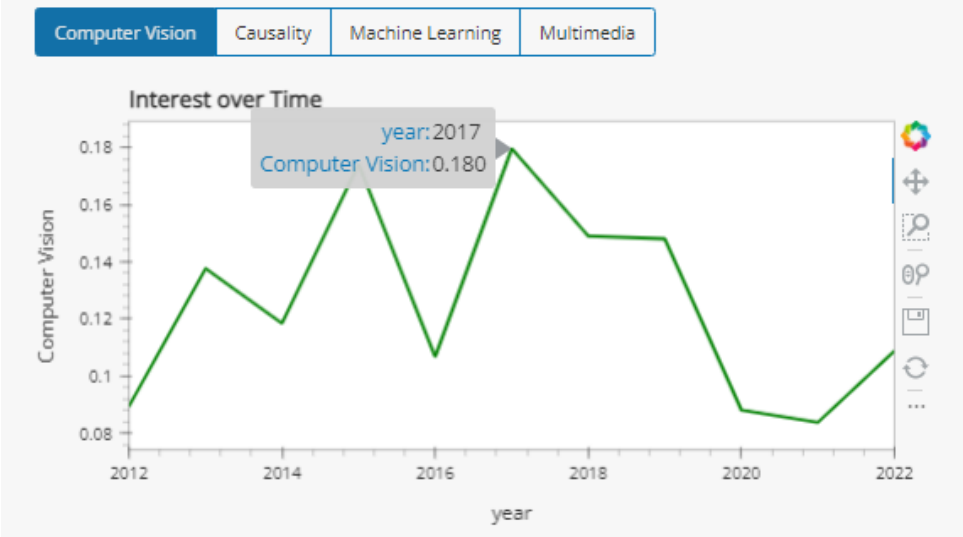


Figure 7- Hanwang Zhang's Interest over time plot - Computer vision

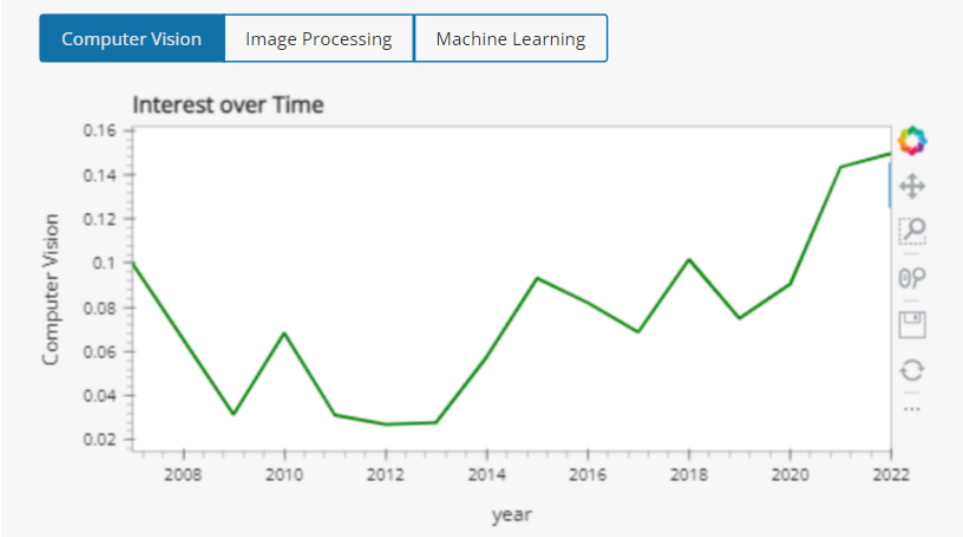


Figure 8 - Chen Change Loy's interest over time plot - computer vision

To acquire such a plot, I computed the semantic similarity between each of the interest tags and every title (and abstract) of the researcher's projects. For instance, given {'Computer Vision': 0.22388215, 'Image Processing': 0.5174386, 'Machine Learning': 0.034008157}, we can see that the researcher was most keen in image processing for this project. Then, for every year compute the mean interest levels towards each interest tag (Figure 9). To compute semantic similarity between the interest tags and the project title and abstract, I used a Python Library, *sentence_transformers.*

	year	Computer Vision	Image Processing	Machine Learning
0	2007	0.099788	0.139104	0.075266
1	2009	0.031384	0.339146	0.023455
2	2010	0.068357	0.329368	0.063951
3	2011	0.031162	0.032334	0.023784
4	2012	0.026979	0.050565	0.190115
5	2013	0.027772	0.037060	0.033313
6	2014	0.057411	0.111568	0.129144
7	2015	0.093192	0.208842	0.170521
8	2016	0.082035	0.127313	0.136565
9	2017	0.068689	0.133076	0.086867
10	2018	0.101666	0.224434	0.167452
11	2019	0.074954	0.146026	0.137883
12	2020	0.090369	0.217872	0.111548
13	2021	0.143467	0.279925	0.171243
14	2022	0.149652	0.213764	0.255738

Figure 9 - Preprocessed data - mean interest score for each interest tag by year (not part of dashboard)

Moving on, to the right of the interest plot, I display the number of citations over time for the SCSE researcher. Like the interest plot, we can hover over the line and observe the exact number of citations the researcher has for that year (Figure 10 and 11).

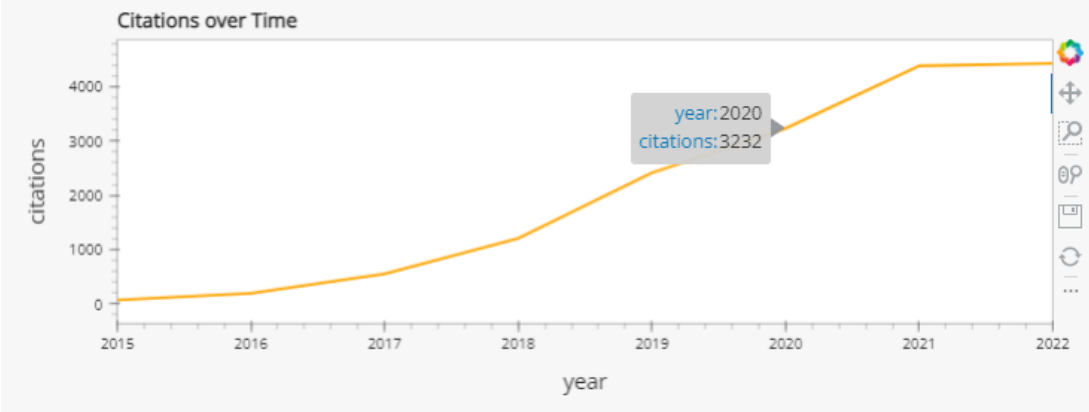


Figure 10 - Citations over time - Hanwang Zhang

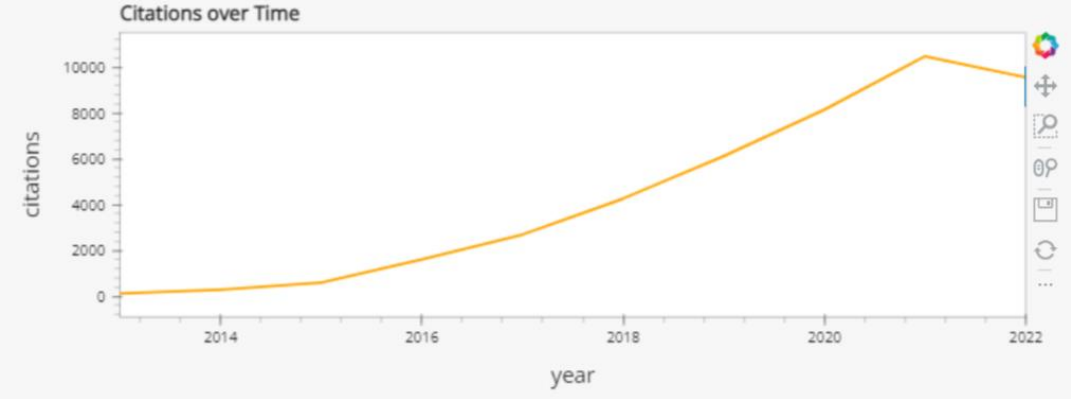


Figure 11 - Citations over time - Chen Change Loy

Right below the interest plot, I display the researchers top 20 co-authors or collaboration. Furthermore, I highlight researchers that were from NTU with a different color. From Figure 12, we can observe that Chen Change Loy has mostly external co-authors. This information was acquired from the researcher's google scholar profile. To create this plot, I used a Python Library, *networkx.*

Figure 12 - Co-author network - Chen Change Loy

To the right of the author's collaboration network, I display a bar plot showing the quality of the researcher's conferences (Figure 13). We can observe that, for Chen Change Loy, most of his projects were published in high quality conferences with A* ranking. For data acquisition, I scraped the conference rankings from resurchify.com, and matched each of the conference to the researcher's projects by the conference acronym.

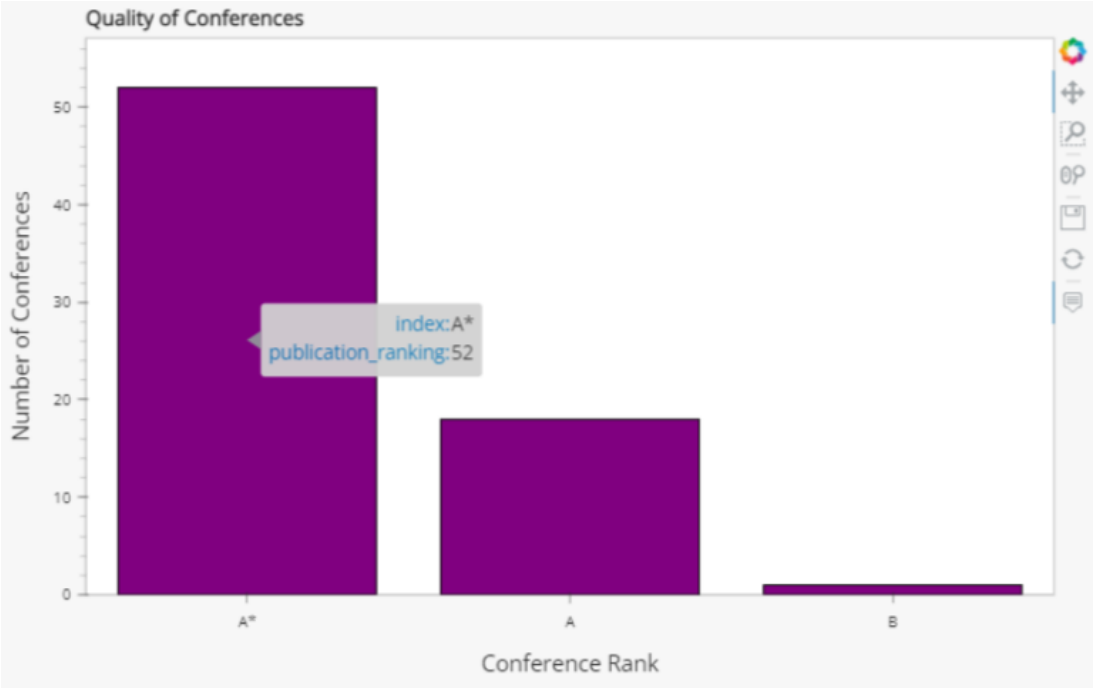
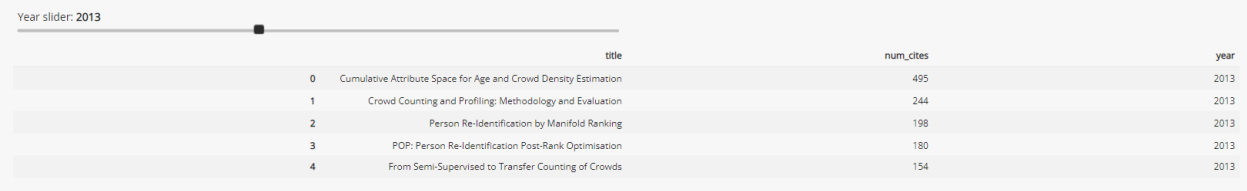


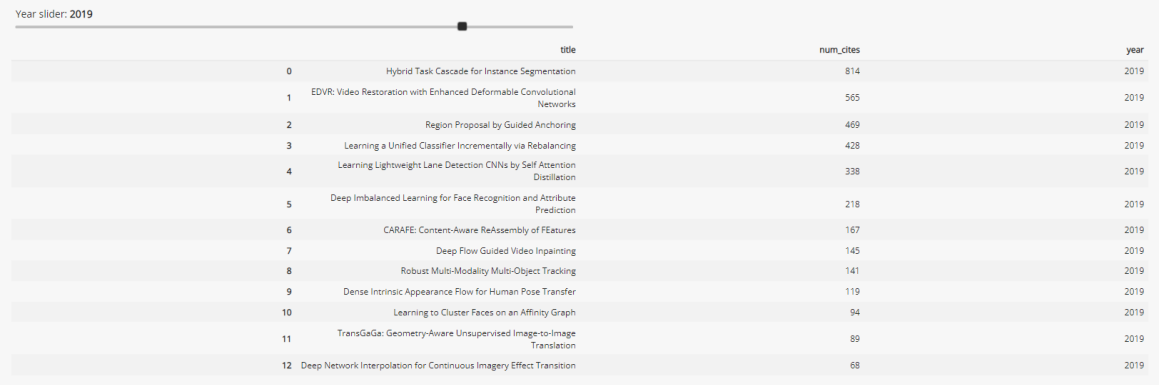
Figure 13 - Publication Conference Rank distribution - Chen Change Loy

Moving further down the dashboard, I display the exact projects completed by the researcher by year. In figure 14, we can observe that Chen Change Loy published 5 research projects in 2013. In 2019, he published 13 research projects (Figure 15). Like the interest plots, we can transition between years or interests by interacting with the scroll bar or buttons. The research projects are sorted by number of citations by default.



	title	num_cites	year
0	Cumulative Attribute Space for Age and Crowd Density Estimation	495	2013
1	Crowd Counting and Profiling: Methodology and Evaluation	244	2013
2	Person Re-Identification by Manifold Ranking	198	2013
3	POP: Person Re-Identification Post-Rank Optimisation	180	2013
4	From Semi-Supervised to Transfer Counting of Crowds	154	2013

Figure 14 - Research Projects by year - Chen Change Loy - 2013

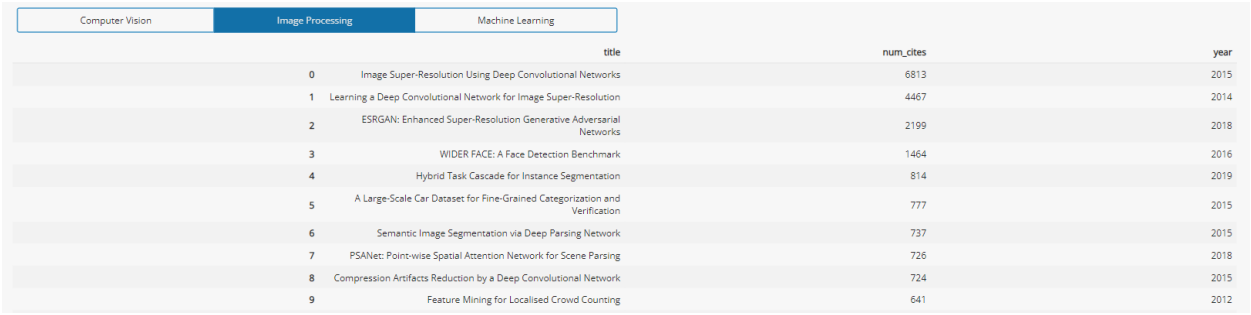


Year slider: **2019**

	title	num_cites	year
0	Hybrid Task Cascade for Instance Segmentation	814	2019
1	EDVR: Video Restoration with Enhanced Deformable Convolutional Networks	565	2019
2	Region Proposal by Guided Anchoring	469	2019
3	Learning a Unified Classifier Incrementally via Rebalancing	428	2019
4	Learning Lightweight Lane Detection CNNs by Self Attention Distillation	338	2019
5	Deep Imbalanced Learning for Face Recognition and Attribute Prediction	218	2019
6	CARAFE: Content-Aware ReAssembly of FEatures	167	2019
7	Deep Flow Guided Video Inpainting	145	2019
8	Robust Multi-Modality Multi-Object Tracking	141	2019
9	Dense Intrinsic Appearance Flow for Human Pose Transfer	119	2019
10	Learning to Cluster Faces on an Affinity Graph	94	2019
11	TransGaGa: Geometry-Aware Unsupervised Image-to-Image Translation	89	2019
12	Deep Network Interpolation for Continuous Imagery Effect Transition	68	2019

Figure 15 - Research Projects by year - Chen Change Loy - 2019

Lastly, I group the research projects by interest and display them in a data frame (Figure 16). The methodology is similar with the previous interest plot. For every research project, I assign it to the most semantically similar interest tag using the *sentence_transformer* library. Different researchers will have different interest tags. For instance, Eric Cambria has the interest tags affective computing, sentiment analysis and multimodal interaction etc. (Figure 17). Research projects are sorted by number of citations by default.



Computer Vision | Image Processing | Machine Learning

	title	num_cites	year
0	Image Super-Resolution Using Deep Convolutional Networks	6813	2015
1	Learning a Deep Convolutional Network for Image Super-Resolution	4467	2014
2	ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks	2199	2018
3	WIDER FACE: A Face Detection Benchmark	1464	2016
4	Hybrid Task Cascade for Instance Segmentation	814	2019
5	A Large-Scale Car Dataset for Fine-Grained Categorization and Verification	777	2015
6	Semantic Image Segmentation via Deep Parsing Network	737	2015
7	PSANet: Point-wise Spatial Attention Network for Scene Parsing	726	2018
8	Compression Artifacts Reduction by a Deep Convolutional Network	724	2015
9	Feature Mining for Localised Crowd Counting	641	2012

Figure 16 - Research Projects by interest - Chen Change Loy

affective computing | sentiment analysis | multimodal interaction | commonsense reasoning | sentic computing

	title	num_cites	year
0	New Avenues in Opinion Mining and Sentiment Analysis	1397	2013
1	SenticNet 5: Discovering Conceptual Primitives for Sentiment Analysis by Means of Context Embeddings	668	2018
2	Context-Dependent Sentiment Analysis in User-Generated Videos	503	2017
3	Targeted Aspect-Based Sentiment Analysis via Embedding Commonsense Knowledge into an Attentive LSTM	499	2018
4	Sentic Patterns: Dependency-Based Rules for Concept-Level Sentiment Analysis	411	2014
5	The Hourglass of Emotions	378	2012
6	DialogueRNN: An Attentive RNN for Emotion Detection in Conversations	353	2019

Figure 17 - Research Projects by interest - Eric Cambria

This concludes all the features for the SCSE researcher dashboard. In summary, I show basic information such as experience and education on the left panel. On the right panel, I display panels showing the number of citations, interest over time, co-author network, quality of conferences, and specific research projects. The data sources were, www.resurchify.com, LinkedIn and Google Scholar. Tools used were Python libraries namely: *Panel, sentence_transformers and networkx.*

SCSE Faculty Profile Dashboard

Next, I will describe the SCSE faculty profile dashboard which aggregates the information from the SCSE researcher profiles. Figures 18 to 19, display the full dashboard.

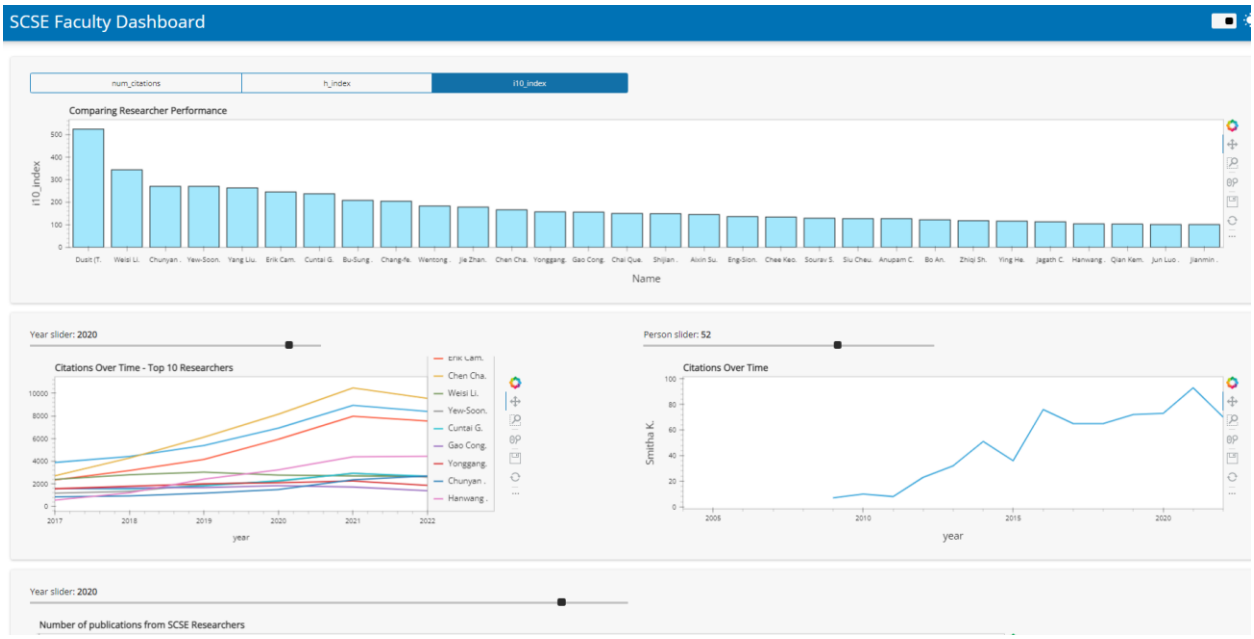


Figure 18 - SCSE Dashboard Part 1

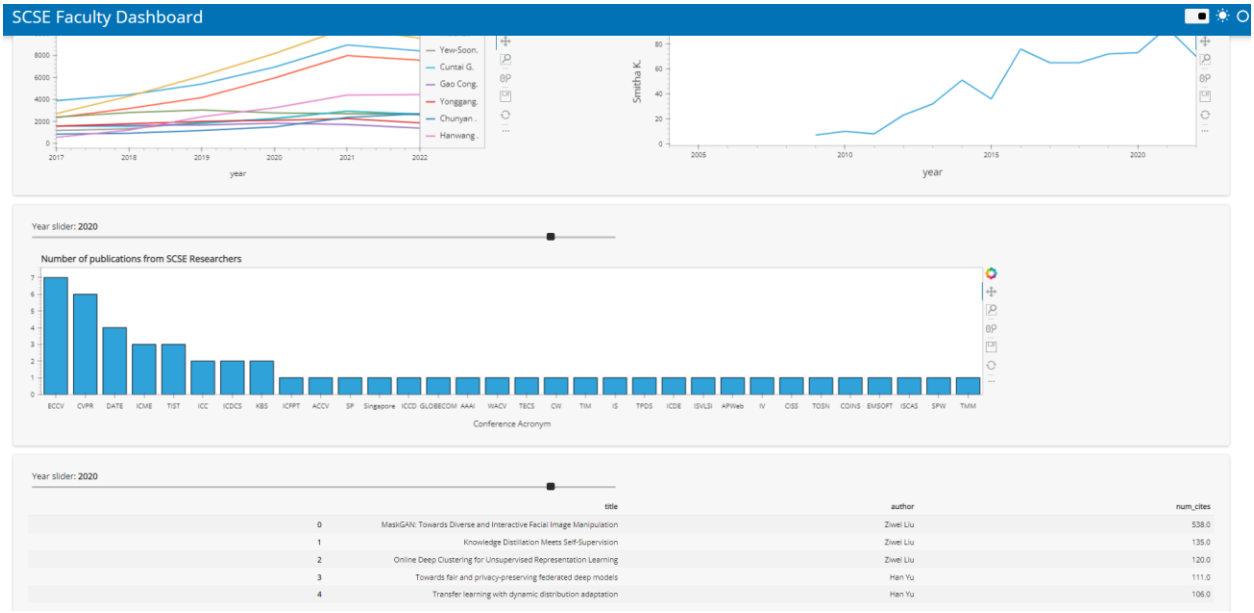


Figure 19 - SCSE Dashboard Part 2

The first panel compares the quality and quantity of research between SCSE researchers (Figures 20 to 22). The metrics compared are number of citations, h-index, and i10-index. From figures 20 to 22, we can observe that Dusit Niyato is ranked one in all three metrics.

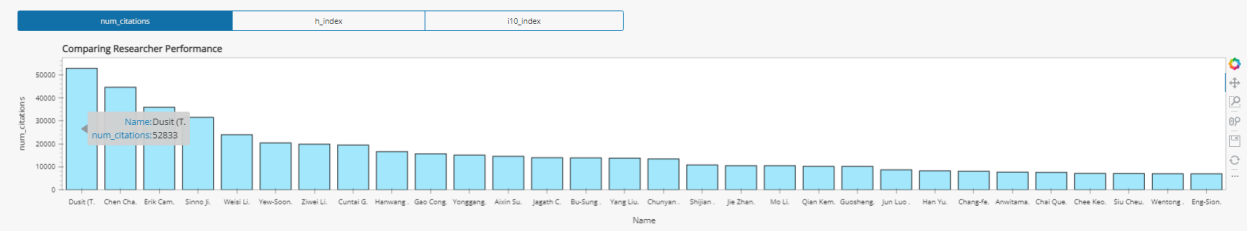


Figure 20 - bar plot comparing number of citations

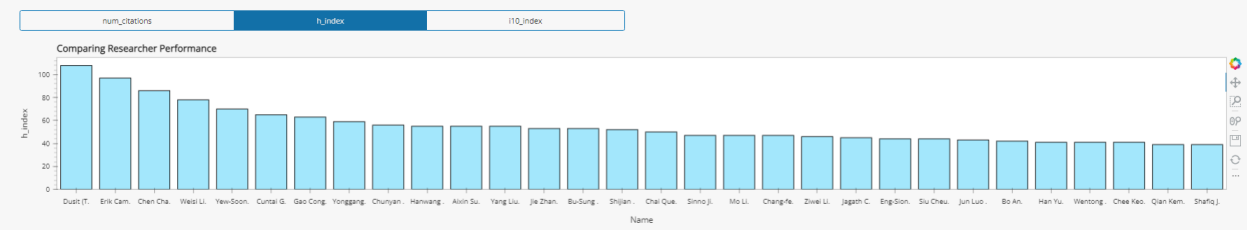


Figure 21 - bar plot comparing h index

Figure 22 - bar plot comparing i10-index

Right below, I compare the citations over time of the top 10 researchers (Figure 23). I limited the number of researchers to 10 to reduce clutter. Furthermore, I implement a year slider too to improve clarity. For instance, in recent years, Chen Change Loy has a relatively fast growth in number of citations (Figure 23), while further back in the past, it was Dusit Niyato with the fastest growth in citations (Figure 24). Furthermore, we can “turn off” certain lines by clicking on the legend, to reduce clutter and improve clarity (Figure 24).

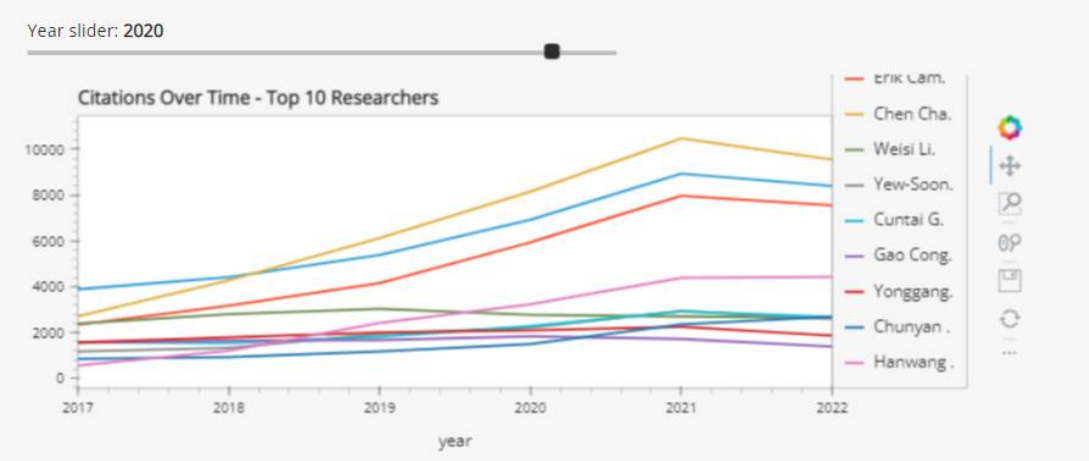


Figure 23 - Line plot comparing growth of citations over time of the top 10 SCSE researchers 1

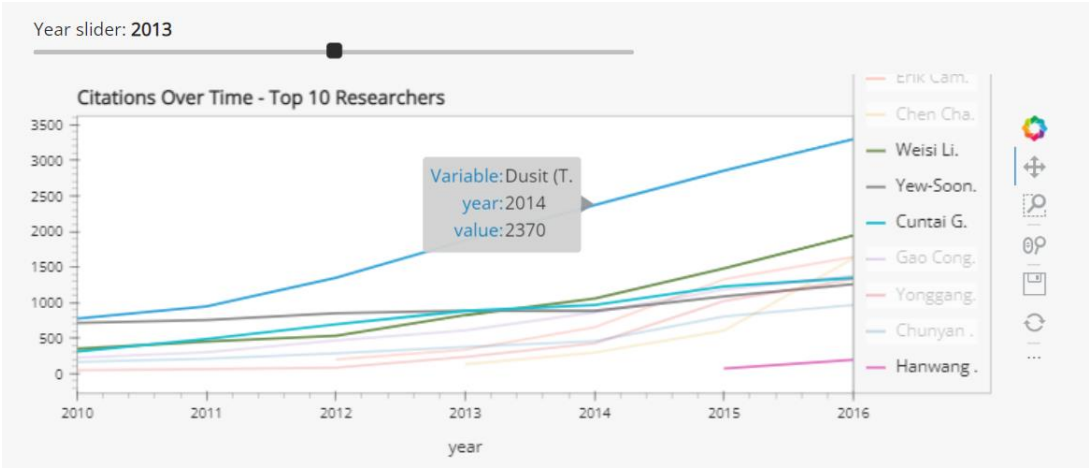


Figure 24 - Line plot comparing growth of citations over time of the top 10 SCSE researchers 2

Since the previous plot, only compare the top 10 researchers (by number of citations), I implemented an alternative line plot that can display the growth of number of citations for any SCSE researcher. To transition between different researchers, simply scroll through the scroll bar (Figure 25 and 26). The researchers are sorted alphabetically to improve usability (Figure 26).

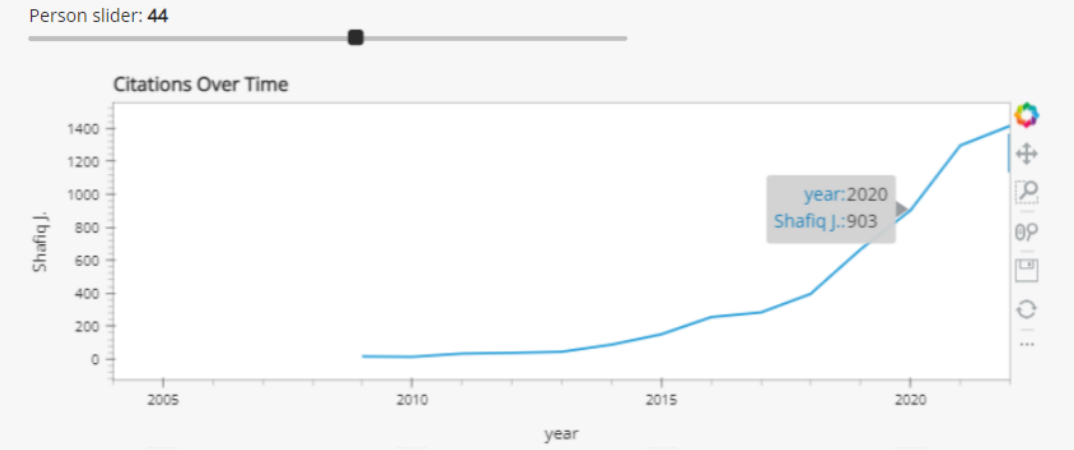


Figure 25 - Interactive line plot that shows citations growth for any SCSE researcher Shafiq J.

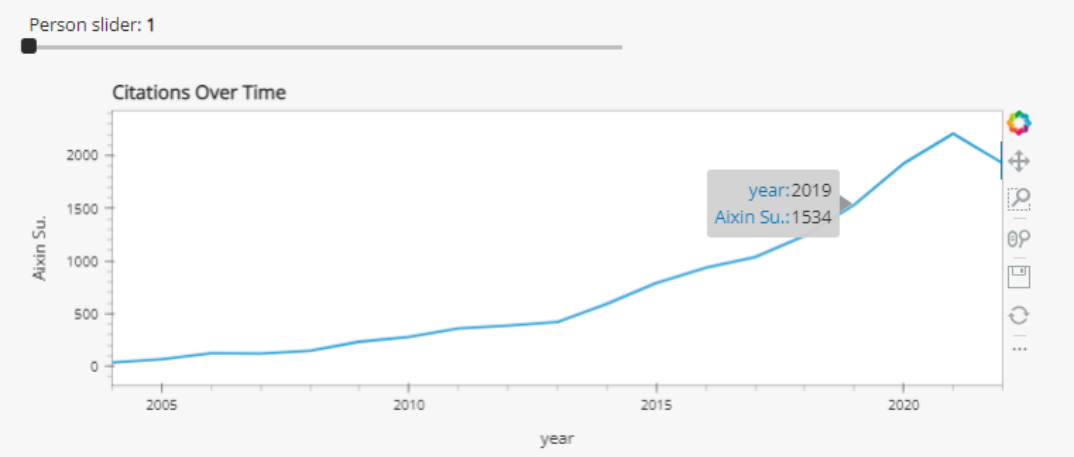


Figure 26 - Interactive line plot that shows citations growth for any SCSE researcher - Sun Aixin

Next, I compare the representation of conferences by SCSE researchers using a bar plot. For instance, Digital Signal Processing (DSP) was the highest represented by SCSE researchers in 2015 (Figure 27). In 2022, Computer Vision and Pattern Recognition (CVPR) was the highest represented by SCSE researchers (Figure 28).

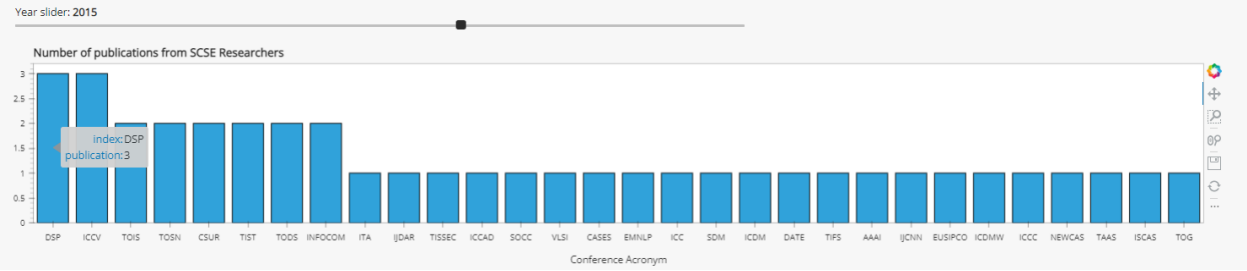


Figure 27 - bar plot comparing the representation of conferences in 2015

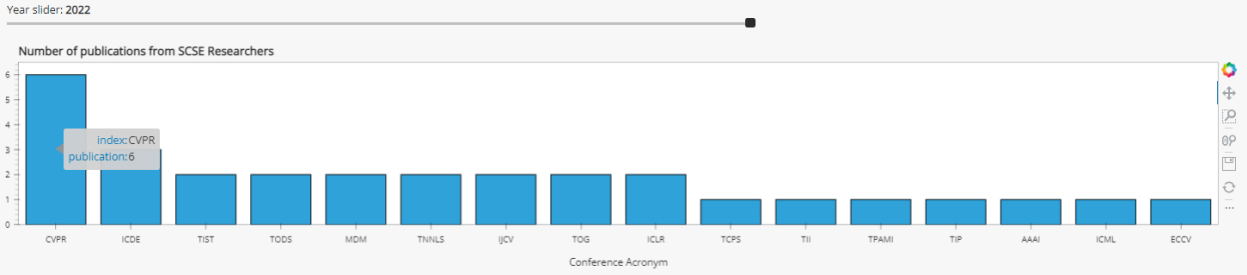
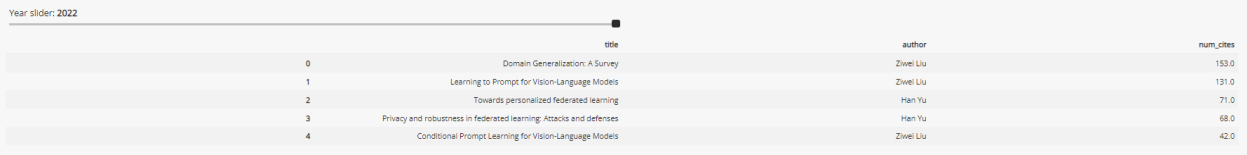


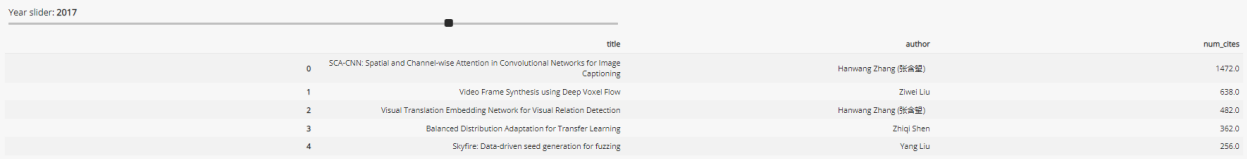
Figure 28 - bar plot comparing the representation of conferences in 2022

Lastly, I implement an interactive data frame that displays the top 5 most cited research projects from SCSE by year. For instance, in 2022, the highest cited paper was “Domain Generalization: A Survey” by Ziwei Liu (Figure 29). In 2017, the highest cited paper was “SCA-CNN:…” by Hanwang Zhang (Figure 30).



	title	author	num_cites
0	Domain Generalization: A Survey	Ziwei Liu	153.0
1	Learning to Prompt for Vision-Language Models	Ziwei Liu	131.0
2	Towards personalized federated learning	Han Yu	71.0
3	Privacy and robustness in federated learning: Attacks and defenses	Han Yu	68.0
4	Conditional Prompt Learning for Vision-Language Models	Ziwei Liu	42.0

Figure 29 - data frame showing the top 5 most cited research papers by SCSE in 2022



	title	author	num_cites
0	SCA-CNN: Spatial and Channel-wise Attention in Convolutional Networks for Image Captioning	Hanwang Zhang (张含望)	1472.0
1	Video Frame Synthesis using Deep Voxel Flow	Ziwei Liu	638.0
2	Visual Translation Embedding Network for Visual Relation Detection	Hanwang Zhang (张含望)	482.0
3	Balanced Distribution Adaptation for Transfer Learning	Zhiqi Shen	362.0
4	Skyfire: Data-driven seed generation for fuzzing	Yang Liu	256.0

Figure 30 - data frame showing the top 5 most cited research papers by SCSE in 2017

In conclusion, the first dashboard displays the profile of any individual SCSE researcher while the second dashboard aggregates the profile information of SCSE researchers into a faculty dashboard. The dashboards created from *panel* are interactive. Multiple tools and data sources are used in the process.