

Ting-Wei Lin¹, Tsun-Wei Tseng², Yu-Hsuan Lin³, Pei-Yu Chen⁴, Ning-Yuan Lyu⁵, Ting-Kuang Lo⁶, Shing-Yun Jung⁷, Hsu Wei⁸, Charles Chuang⁹, Chun-Yu Yin⁸, Johnson Hsieh¹⁰

1.Genome and Systems Biology Degree Program, National Taiwan University and Academia Sinica 2.TAO Info Co.Ltd 3.inQ Technology Co.Ltd 4.Pegatron Co.Ltd 5.Department of Electrical Engineering, National Tsing-Hua University 6.Department of applied mathematics, Feng Chia University 7.Institute of Computer Science and Engineering, National Chiao Tung University 8.Hfoundation 9.NETivism Co.Ltd 10.Department of Computer Science, National Chengchi University

Background

Assessing the troubled families risk status and distributing the resources appropriately is a big issue for nonprofit organizations, not to mention the national program like UK Trouble Families Program¹. Unfortunately, those organizations and national programs still use the conventional way to deal with these problems².

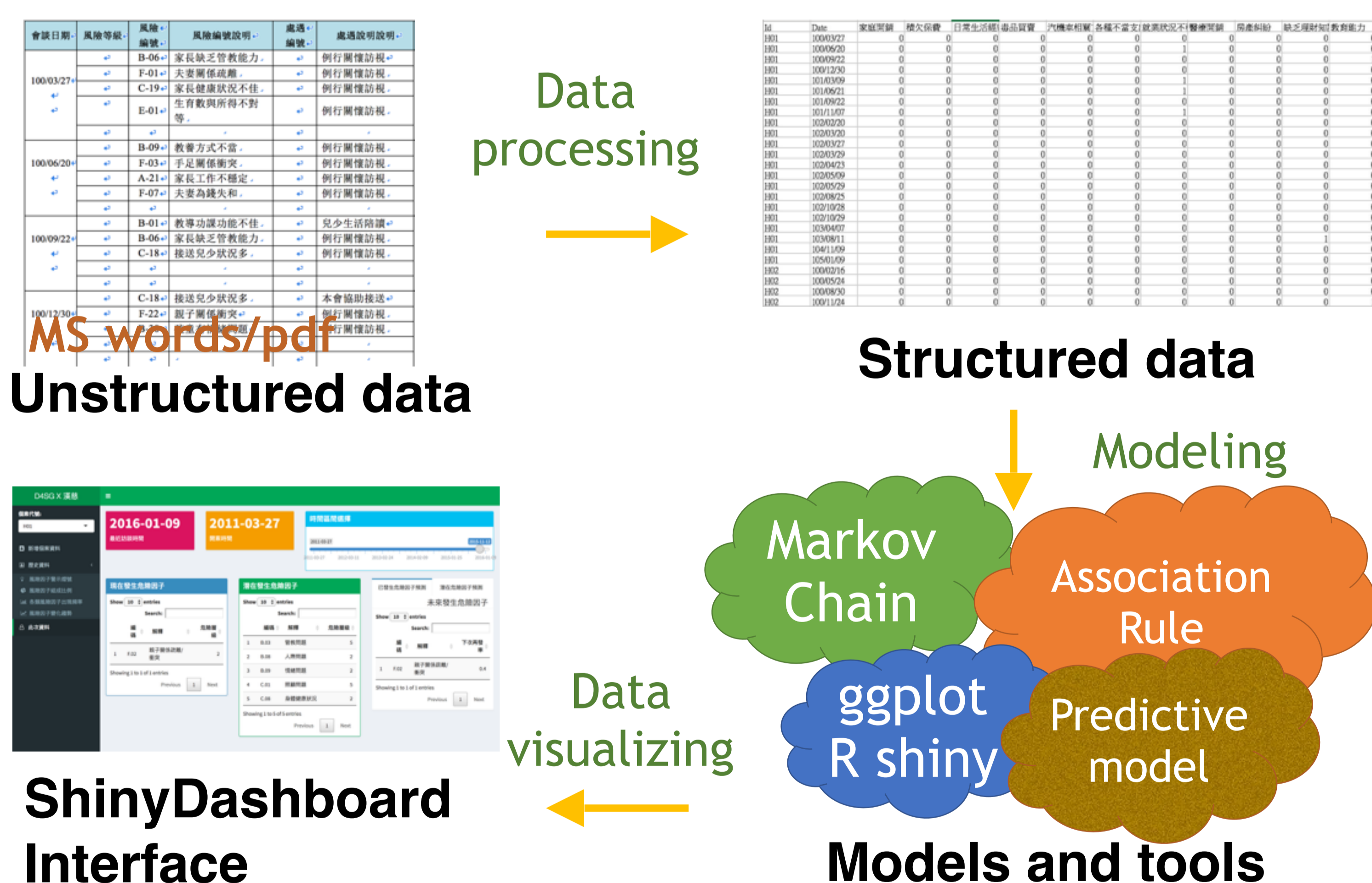
As the increasing demands for social assistance and long-term shortage of social workers in these field, a more precise and continuous way to handle the social resources wisely and efficiently is needed. Besides, the junior social workers are hard to quickly get a hang of several families through lots of interview records and past documents, then decide whether providing their intervention. And most importantly of all, they lack of ability to utilize those information with proper data engineering and summarize those experience through data analysis.

So here We first apply a evidence-based approach on assessing the troubled families' risk status with a R dashboard application integrated the prediction model generated from those families archives, follow-up records under the Data Science For Social Good Program in Taiwan with the cooperation between a volunteer data science team and local nonprofit organization HFoundation.

Objective

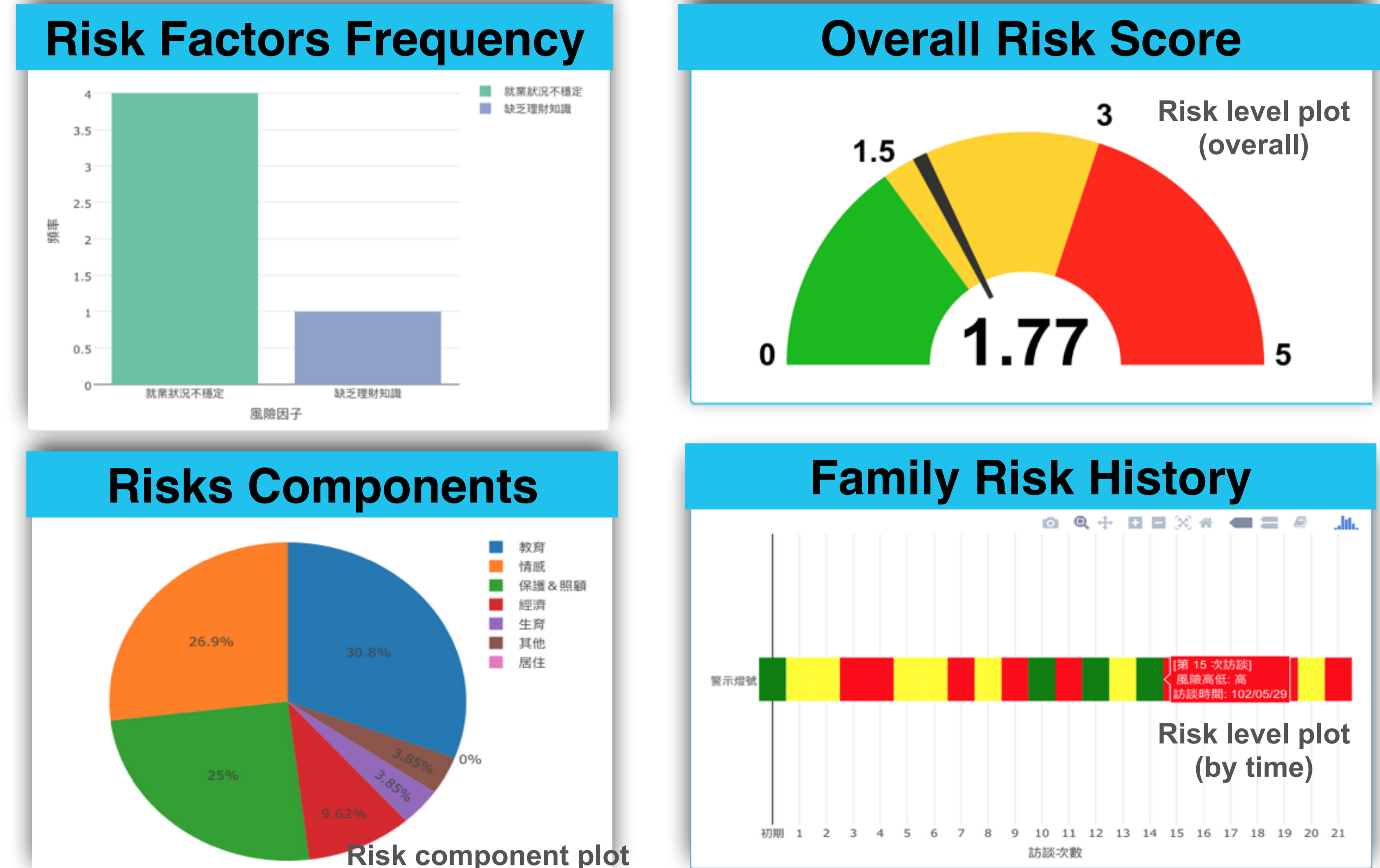
1. Visualize these historical data with the simple and clear plots to help the social workers understand the conditions of different troubled families in a short time
2. Construct the forecasting models to provide information for the social workers to take correct actions to assist the troubled families.
3. Build an electronic system to store all the data of the troubled families. At the same time, social works can also easily create the new data for the new-coming troubled family.

Data process workflow



Trouble Family Monitor Framework

Family Risk factor monitoring



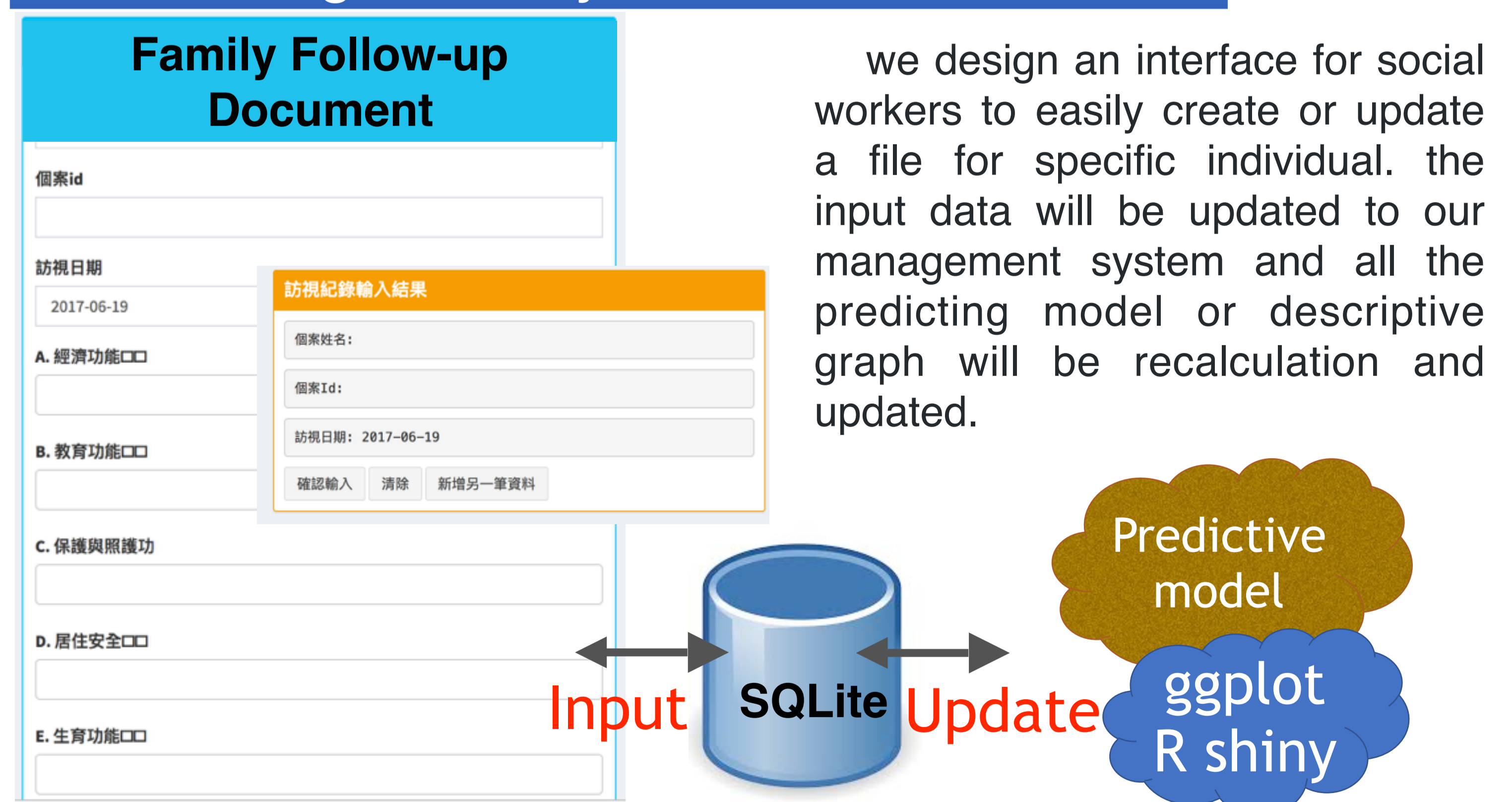
Here are four primary plots created with 'ggplot2', respectively pie charts and histogram to show the composition of risk factors in the family, trend plot to follow-up the variation of risk level by time, speedometer style plot in order to indicate whether the family need more attention.

Prediction of family problems

Current Risk Factors				Potential Risk Factors			
Risk Code	Description	Level	Risk Code	Description	Level		
1	B.06 Sexual Behavior Problem	5	1	A.07 Unemployment	2		
2	B.10 School Learning Problem	1	2	B.03 Child Education Problem	5		
3	F.02 Parent-Child Relationship	2	3	B.08 Social Relationship Problem	2		
			4	B.09 Parent Emotion Problem	2		
			5	C.01 Daily Caring Problem	5		
			6	C.03 Other issue	1		
			7	C.05 Drug addition, alcohol abuse	5		
			8	C.08 Physical Health Problem	2		

we use the **association rule analysis** to reveal the possible unnoticed problems according to family's current risk facts, and construct a **Markov chain model** from past follow up documents to predict future risk factors

Data Management system



Reference

- 1.Fletcher A, Gardner F, McKee M, Bonell C.(2012). The British government's Troubled Families Programme. BMJ2012;344:e3403. doi:10.1136/bmj.e3403 pmid:22592850.
- 2.Chris Bonell, Martin McKee.(2016). Adam Fletcher. Troubled families, troubled policy making. BMJ 2016; 355 doi: https://doi.org/10.1136/bmj.i5879

