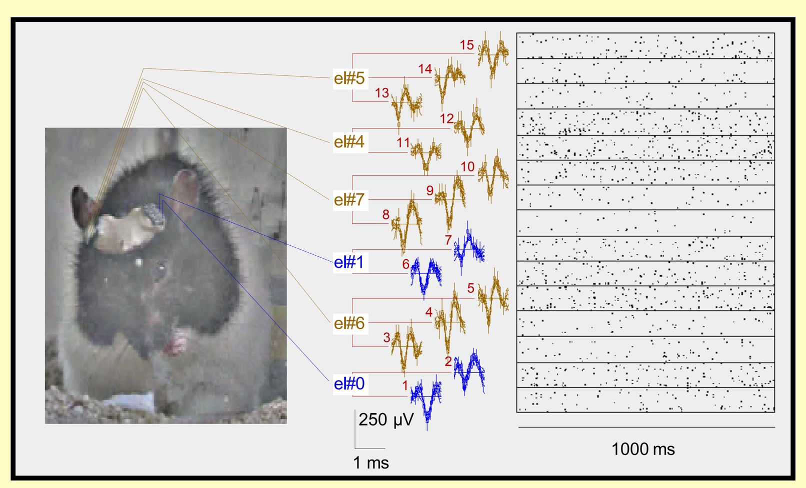
DATA ANALYSIS FOR THE NEUROSCIENCE: A VIRTUAL LABORATORY FOR SCIENTIFIC INVESTIGATION

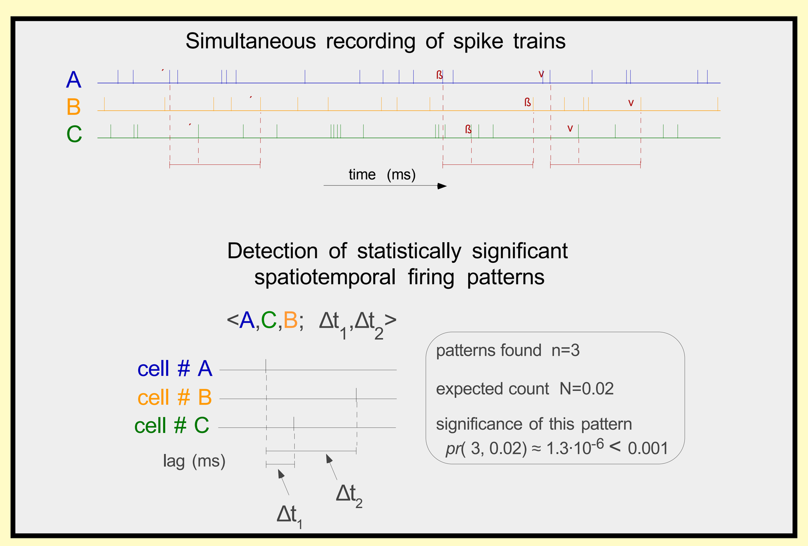
Javier Iglesias[1,2,3], Tetyana Aksenova[2], Volodymyr Volkovitch[4], Alessandro E.P. Villa[1,2]

(1) Inserm U318 & Laboratoire de Neurobiophysique, Université Joseph-Fourier Grenoble, Grenoble, France; (2) Laboratoire de Neuroheuristique, Institut de Physiologie, Université de Lausanne, Switzerland; (3) Computer Science Institute, Collège Propédeutique, Université de Lausanne, Switzerland; (4) Dpt. Control System, International Research Center Information Technologies and Systems, Kiev, Ukraine

why?

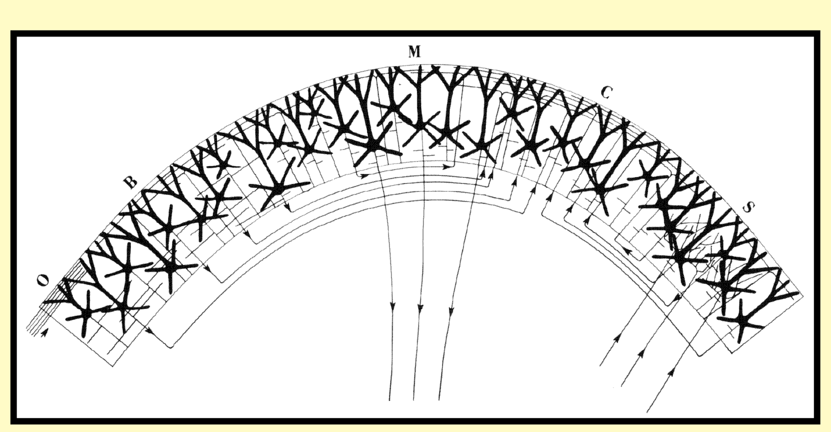
The collective contribution of brain researchers from all over the world can achieve unprecedented results in deciphering neural coding. The sequences of cell discharges--the spike trains--form time series whose dynamics are strongly related to the information processing carried out in the brain areas recorded by extracellular electrophysiological techniques. Most of neural coding theories relate the neurophysiological activity, measured by one or multiple spike trains, with sensory and/or motor events. The study of cell assemblies and the availability of cheap computing power have raised the possibility to analyze the same data set within several theoretical frameworks. Our objective is to develop a modular computing approach, providing a flexible tool that can manage a wide set of analytical methods contributed by investigators connected to Internet.



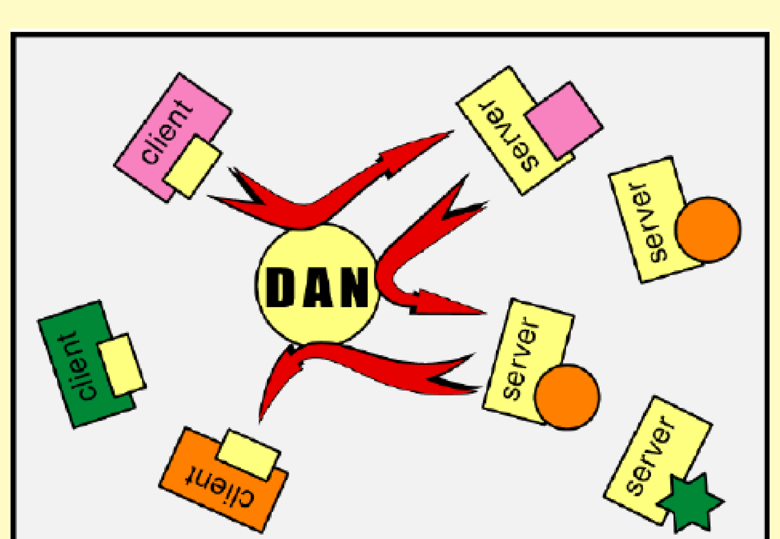


how?

Neural networks are characterized by distributed information processing where local processing power can not achieve the whole task, and cooperation of several units is necessary. In a similar way, we believe that it is important to allow local intelligence, in terms of specific analytical methods, to interact with a distributed amount of data in order to let emerge new and unexpected results and ideas. We propose a software implementation, called DAN (Data Analysis for Neurophysiology or Data Analysis Network), where computer nodes independently process the data stream and communicate final or partial results to any other node for additional processing, storage, or graphical representation. Technically speaking, this is achieved through a combination of several platform-independent Java technologies assembled on top of Internet (Villa et al. 2000, 2001).

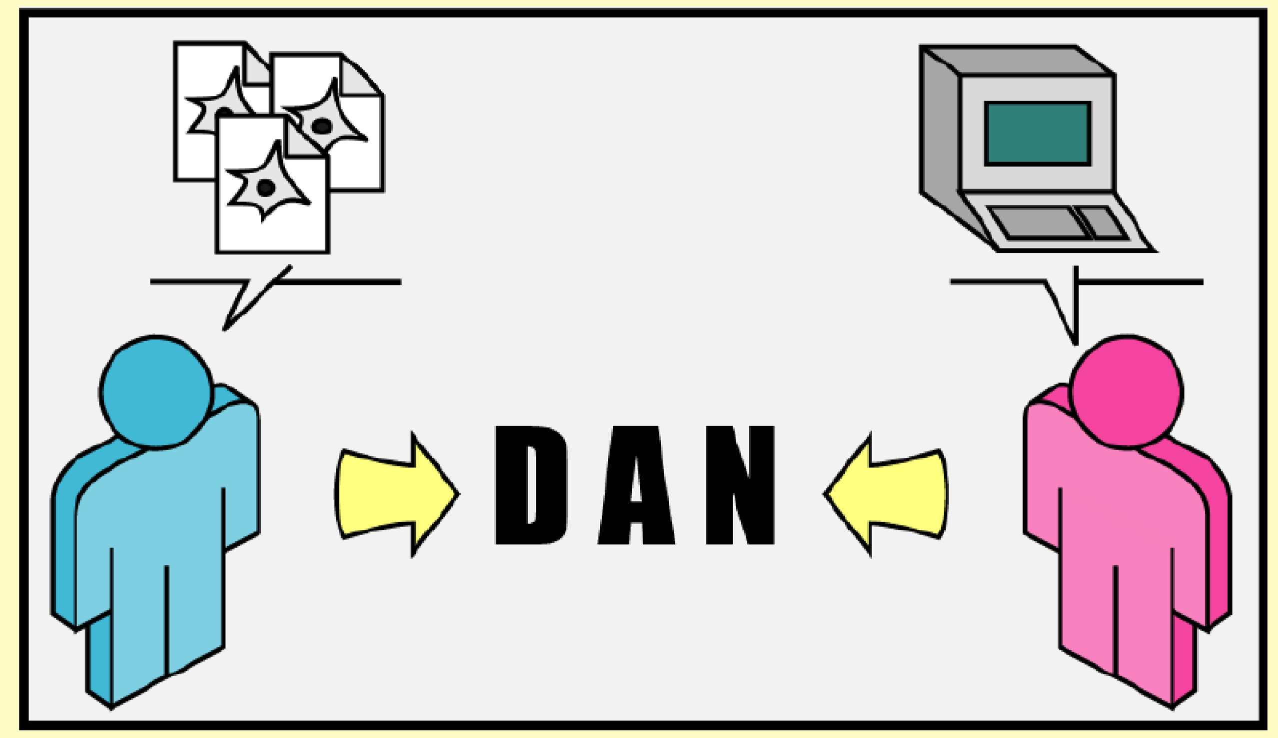


The skeleton cortex (fig. 8.3) from "On textures of Brain", by Valentino Braitenberg, Springer Verlag, 1977.



http://dan.unil.ch

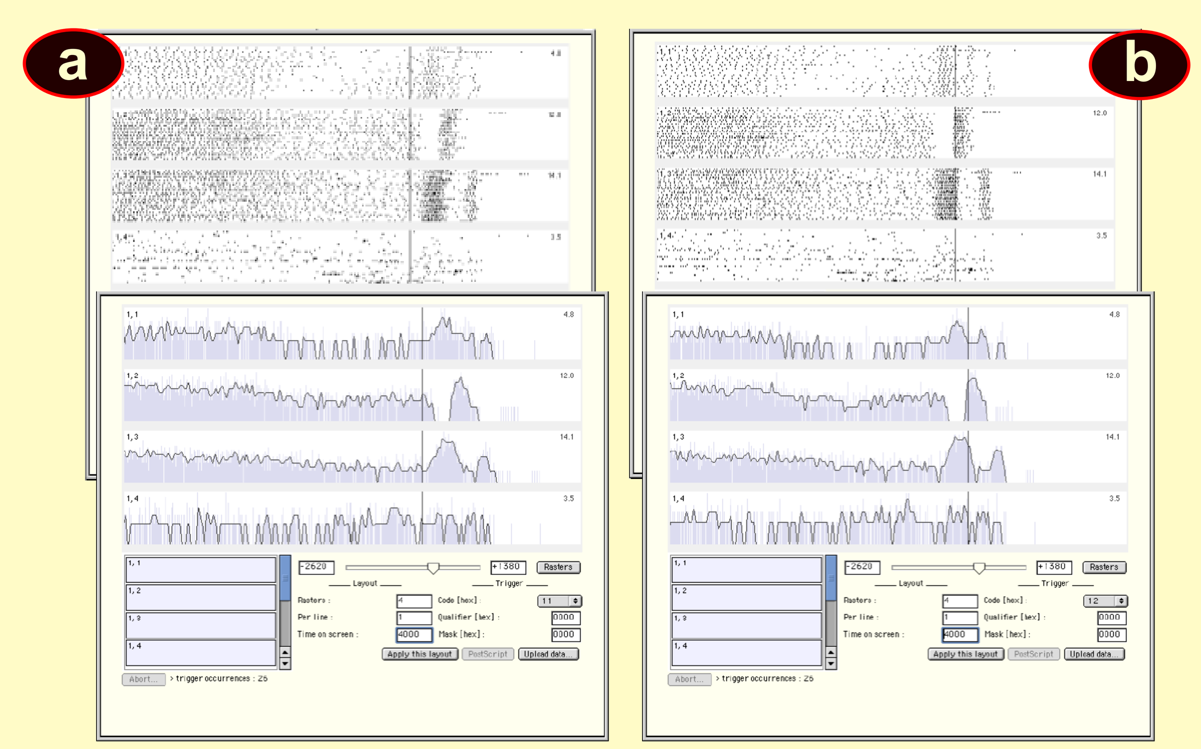
to whom?



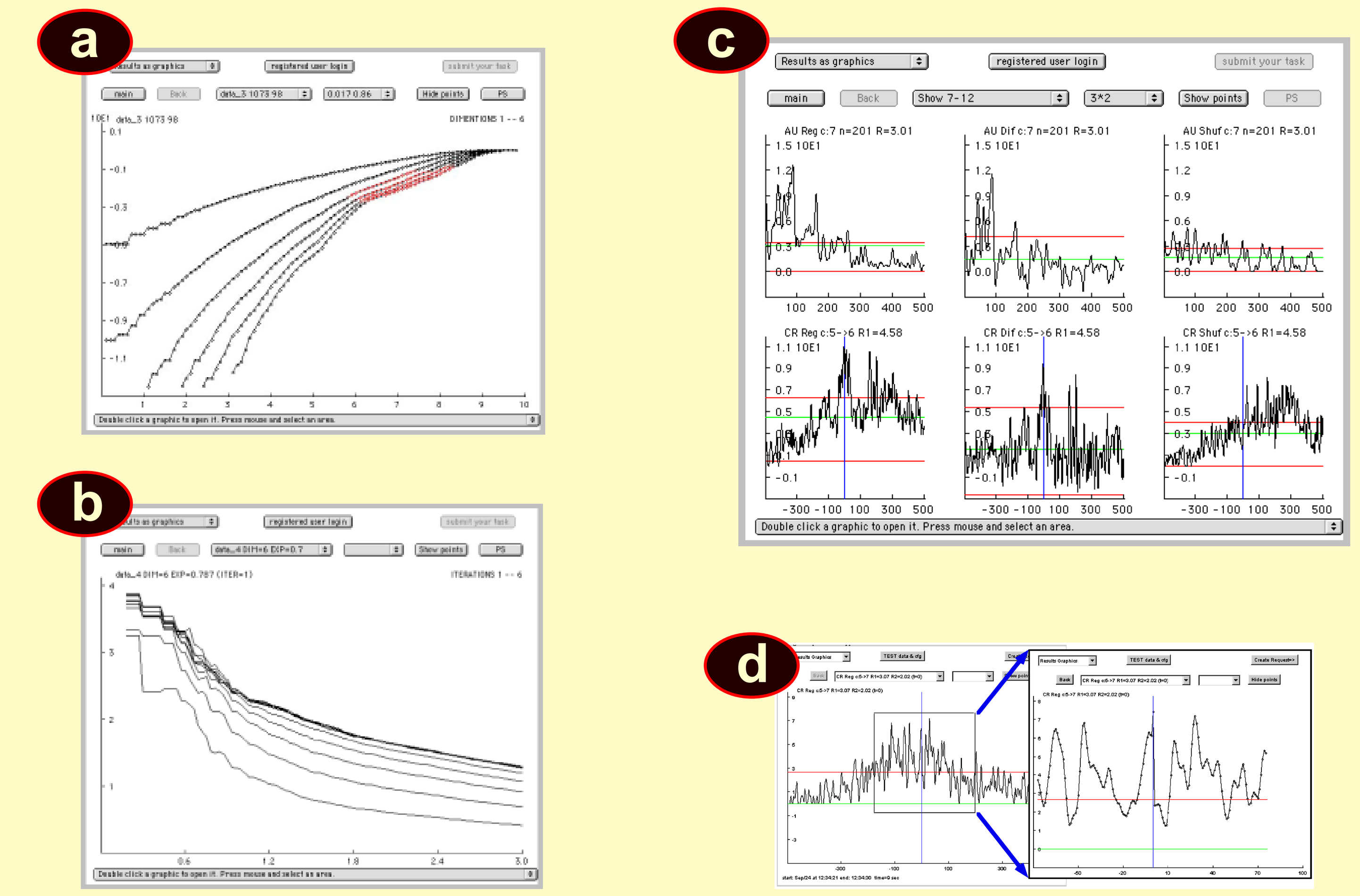
There are two types of users. The first type is composed of users who develop new theories and analytical methods. They would like to offer and spread the result of their effort throughout the scientific community of neuroscience researchers. It often happens that this information remains in the literature for a long time before practical applications, if any, become available to a wide range of experimentalists. The second type is formed by users who collect lots of data. They would like to test several hypotheses and related analytical methods in order to understand the message carried out by their recorded data. Among the users of this category we find the people who collect data from the living brain and from simulated neuronal networks (*http://www.u444.jussieu.fr/xnbc*). Both types of users are willing to establish a correlation between theory and experiments, but they often lack the appropriate environment.

when?

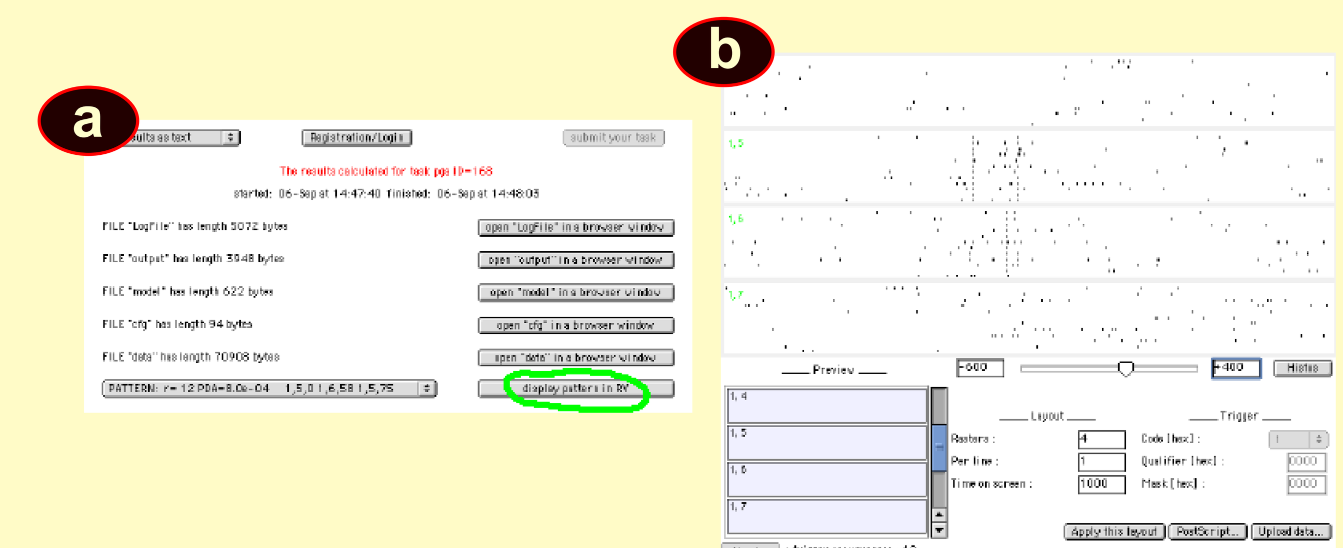
At present time a limited number of time series analyses (auto- and cross-correlations, Grassberger-Procaccia correlation dimension analysis, maximum Lyapunov exponent, complex search for spatio-temporal patterns) and graphical representation of time series (referred to as RasterViewer) are freely available through the web portal of the Laboratory of Neuroheuristic (http://www.neuroheuristic.org).



RasterViewer: dot rasters of multiple spike train data and peri-event time histograms of motor cortex single unit activities recorded in a simple reaction time task (data kindly provided by Riehle et al., CRNC-CNRS, Marseille). The data can be triggered with the stimulus onset (**a**) or with the movement onset (**b**).



a - Lines: Embedding and correlation dimensions using the Grassberger-Procaccia algorithm are calculated and the scaling region is determined automatically.
b - Lines: The largest Lyapunov exponent allows to quantify the divergence of nearby trajectories in the data sets as a function of the embedding space.
c - Histograms: Several renewal density histograms of auto- and cross-correlations can be displayed on the same page
d - Histograms: same as (c) displayed as a single graph and zoomed by the user.



a - Output frame of the results obtained after search of spatiotemporal patterns. The selected pattern (lower left) may be displayed immediately in the raster form.
b - Raster display of the activity of four single units aligned by displaying the first spike in pattern **<5,6,5 ; 58,133>** at time 0.

In particular, complex spatiotemporal patterns of spikes can be displayed with many options as a function of significant events.

so what?

We believe that the solitary effort to decipher the neural code represents a dead end. Collaboration between theoreticians and skilled experimentalists is crucial and we expect the number of available methods and experimentalist users to increase with the help of interested people met at this workshop.

References

Villa A.E.P., Tetko I.V., Iglesias J. (2001) Computer Assisted Neurophysiological Analysis of Cell Assemblies, *Neurocomputing* 38-40: 1025-1030.11.
Villa A.E.P., Tetko I.V., Iglesias J., Filipov D. (2000) Transdisciplinary approach to scientific data analysis through Internet, in : *Transdisciplinarity: Joint Problem-Solving among Science, Technology and Society* (R. Häberli, R.W. Scholz, A. Bill, M. Welti Eds.) Haffmans Sachbuch Verlag/Zürich, pp. 550-555.