

**Master Traitement Automatique des Langues
(TAL)
M.Sc. Natural Language Processing (NLP)**

September 7th, 2020

Welcome!

Covid-19

Current guidelines:

- everyone should wear a mask all the time,
- a one way traffic is set up in the building
- the physical distance must be maintained.
- room capacities are adapted (thanks to hard work from colleagues)
- wash your hand frequently and use hand sanitizer

Current guidelines:

- everyone should wear a mask all the time,
- a one way traffic is set up in the building
- the physical distance must be maintained.
- room capacities are adapted (thanks to hard work from colleagues)
- wash your hand frequently and use hand sanitizer

- do not come if you have symptoms
- if you have COVID, you must inform the administration:
`medecine-etudiante-coronavirus@univ-lorraine.fr`

Covid-19 Crisis

When you enter a room, fill the line **starting at the back** and the squares closest to a wall

Use **hydroalcoholic gel** (dispenser will be installed in the corridor)

Rooms will be disinfected **during the lunch break** and every evening

The ventilation of the rooms is done with **fresh air**

Covid-19 Crisis

When you enter a room, fill the line **starting at the back** and the squares closest to a wall

Use **hydroalcoholic gel** (dispenser will be installed in the corridor)

Rooms will be disinfected **during the lunch break** and every evening

The ventilation of the rooms is done with **fresh air**

You have mask in your **welcome bag**

Respect the rules

in case of cluster, **the master/IDMC can operate remotely**

Most of de courses are presential one but some can be remotely given.

Most of the courses are presentational one but some can be remotely given.

Note that the **health authorities** can decide to lockdown a class, a master, the building, at any time in view of the health situation.

Most of the courses are presential one but some can be remotely given.

Note that the **health authorities** can decide to lockdown a class, a master, the building, at any time in view of the health situation.

So for now, courses are **presential**.

Most of the courses are presential one but some can be remotely given.

Note that the **health authorities** can decide to lockdown a class, a master, the building, at any time in view of the health situation.

So for now, courses are **presential**.

In any case, keep us posted.

FIRST!

Registration! = ENT, schedule, mailing list, Crous, ...

You need **your laptop!**

(if you don't have, the IDMC can lend one (we have a limited number of computer, so, go to ask for one now))

contact: `laura.muratore@univ-lorraine.fr`

Join the Discord Server!

(contact me if needed `MaximeAmblard#3475`)

NLP master

- Computer Science on language
- use of the AI's tools and methods

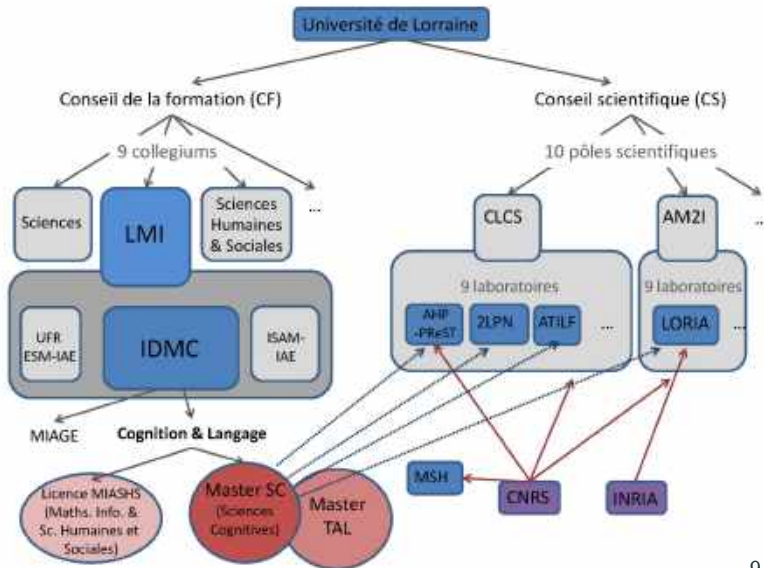
- Computer Science on language
- use of the AI's tools and methods

- Where:
 - new technology companies with a strong innovation component,
 - R&D
 - academic research

- What:
 - define and implement tools and services using language data
 - speech processing, data parsing and understanding, knowledge



- Université de Lorraine (UL)
Pierre mutzenhardt
- Collegium Lorraine Management
Innovation (LMI)
Elisabeth Deschanet
- Institut des Sciences du Digital,
Management et Cognition (IDMC)
Antoine Tabbone
- Master TAL / NLP
Maxime Amblard



Organizational staff

- **Maxime Amblard**, head of the Master TAL,
`maxime.amblard@univ-lorraine.fr`, LORIA B226
 - in charge of the master
 - in charge of the master 1
 - office hour: Monday 12:30-13:30
- **Miguel Couceiro**, Erasmus Mundus LCT local coordinator,
`miguel.couceiro@univ-lorraine.fr`, LORIA B-144
 - in charge of the master 2
 - office hour: Monday 12:00-13:00 (subject to adjustments)
- **Karine Weisse**, administrative issues,
`karine.weisse@univ-lorraine.fr`,
- **Marie-Luce Boulet**,
`marie-luce.boulet@univ-lorraine.fr`,

- DESS TEXTE

(Fiammetta NAMER and Yannick TOUSSAINT)

- Master SCA - Sciences de la Cognition et ses Applications

(Abdel Belaid, Maxime Amblard, Manuel Rebuschi)

- elective courses in M1
- specific year in M2

(Guy Perrier, Fabienne Venant, Maxime Amblard)

- Erasmus Mundus LCT

(Patrick Blackburn, Maxime Amblard, Miguel Couceiro)

- Master SDL and master Computer Science

- URL: <http://institut-sciences-digitales.fr/>
- Mailing list
idmc-m1-tal-2021@etu.univ-lorraine.fr
idmc-m2-tal-2021@etu.univ-lorraine.fr
- Schedule (ADE): <https://ent.univ-lorraine.fr/>
(many updates!)
- Course platform: <https://arche.univ-lorraine.fr>
- Administrative registration:
<https://inscriptions.univ-lorraine.fr>
- Pedagogical registration (IP) <https://ipweb.univ-lorraine.fr>
- grades etc. <https://ent.univ-lorraine.fr>

We welcome students from:

- Students of the master
- in-company student
(contract can be signed during September)

We welcome students from:

- Students of the master
- in-company student
(contract can be signed during September)
- master SDL (for M1 courses in prog. corpus and logic)
- École des Mines (not this year)
- LCT Erasmus Mundus (that's student of the master in fact)



- EMLEX (sometimes)

We welcome students from:

- Students of the master
- in-company student
(contract can be signed during September)
- master SDL (for M1 courses in prog. corpus and logic)
- École des Mines (not this year)
- LCT Erasmus Mundus (that's student of the master in fact)



- EMLEX (sometimes)

and we recommend to discuss with everyone

We welcome students from:

- Students of the master
- in-company student
(contract can be signed during September)
- master SDL (for M1 courses in prog. corpus and logic)
- École des Mines (not this year)
- LCT Erasmus Mundus (that's student of the master in fact)



- EMLEX (sometimes)

and we recommend to discuss with everyone

(keeping social distancing!)

Accords Erasmus



- Trento (Italie)
- Dublin (Irlande)
- Osnabruck (Allemagne)
- Bucarest (Roumanie)
- Malte (Malte)

Accords Erasmus



- Trento (Italie)
- Dublin (Irlande)
- Osnabruck (Allemagne)
- Bucarest (Roumanie)
- Malte (Malte)
- Only **second semester** of first year (application 09.15)

Accords Erasmus



- Trento (Italie)
- Dublin (Irlande)
- Osnabruck (Allemagne)
- Bucarest (Roumanie)
- Malte (Malte)
- Only **second semester** of first year (application 09.15)
- M2's internship can be done abroad
- contact :
laurent.thomann@univ-lorraine.fr

Links with Research and Companies

- lot of things in courses for both!
- **projects** in courses (see this afternoon with me)

Links with Research and Companies

- lot of things in courses for both!
- **projects** in courses (see this afternoon with me)
- **Supervised Project** (see details after): one year work on ho research topic, supervised by a researcher.
- **Seminar** and **Forum** (not this year)
- **3-months internship** between first and second year
- **Software project**
- 4-6-months **internship** during the second year

Links with Research and Companies

- lot of things in courses for both!
- **projects** in courses (see this afternoon with me)
- **Supervised Project** (see details after): one year work on ho research topic, supervised by a researcher.
- **Seminar** and **Forum** (not this year)
- **3-months internship** between first and second year
- **Software project**
- 4-6-months **internship** during the second year
- Alternance

- **LORIA** - Laboratoire lorrain de recherche en informatique et ses applications
UMR 7503, Université de Lorraine, Inria Nancy Grand Est, CNRS
- **Inria - Nancy Grand-Est**
- **ATILF** - Analyse et traitement informatique de la langue française
UMR 7118, Université de Lorraine, CNRS
- **MSH** - Maison des Sciences de l'Homme Lorraine
USR 3216, Université de Lorraine, CNRS
- **Projet Impact LUE - OLKI**

Through its 29 research teams, the LORIA has recognized expertise in Information and Communication Sciences and Technologies. Joint Lab with [Inria](#) and [CNRS](#).

1. Algorithms, Computation, Image and Geometry
2. Formal methods
3. Networks, Systems and Services
4. [Knowledge and Language Management](#)
5. Complex Systems, Artificial Intelligence and Robotics

Knowledge and Language Management

Head: Bruno Guillaume

- **CELLO**: Computational Epistemic Logic
- **équipe K**: Knowledge: Reasoning, Engineering, Data Science
- **ORPAILLEUR**: Knowledge Discovery guided by Domain Knowledge
- **MULTISPEECH**: Speech Modeling for Facilitating Oral-Based Communication
- **READ**: Recognition of writing and analysis of documents
- **SEMAGRAMME**: logic-based models, methods, and tools for the semantic and discourses.
- **SMarT**: Statistical Machine Translation and Speech Modelization and Text
- **SYNALP**: Natural Language Processing : representations, inference and semantics

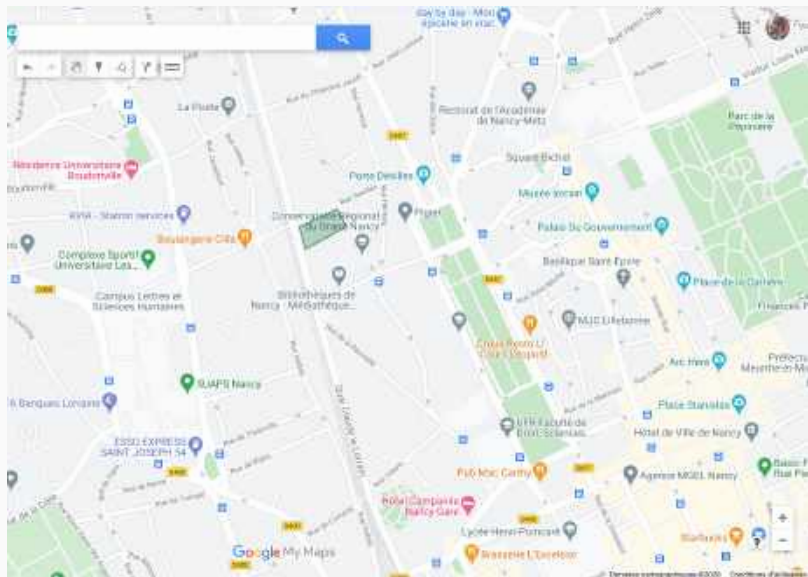
Linguistic Laboratory, a joint research with **CNRS**.

- Linguistique historique française et romane
- Lexique
- Didactique des langues et sociolinguistique (Crapel)
- Discours
- Ressources, normalisation, annotation et exploitation

Where are we



Food, Crous, Break, Library, buy a bike, having a life, ...



(Some) Wednesday afternoon:

- TALC seminar – Loria
- Master seminar – PHS

(Some) Wednesday afternoon:

- TALC seminar – Loria
- Master seminar – PHS

Due to Covid-Crisis, for now, the lab cannot welcome students without contract.

General informations

General informations

- There is a specific course on Arche for each year where you can find links to ressources for the different classes.
 - M1 TAL - NLP
 - M2 TAL - NLP
- For each courses, ask explicitly the evaluation of the courses at the beginning of all courses
- B2E points are not accepted
- Remediation courses (CS and ling.)

Some advices

- Start reading scientific papers,
- Learn Python basics and libraries
- Mathematics are a core component of NLP
- create isolated environment on your computer for NLP projects

Organisation of the semester

Alternance

- The master is open to Alternance. Some students can spend time in companies.
- M1 and M2 have a desynchronise calendar of alternance. Most of the time only 1 year have class.
- The other year can use the other room when it is available
 - keep social distance
 - only 1 student / chair each half day
 - courses of other classes of IDMC have priority
- during the alternance weeks, student can have few courses and must work on project.

Organisation 1st semester M1

The semester is organised in bloc courses:

The image displays 20 individual course timetables for the 1st semester M1, arranged in a 5x4 grid. Each timetable is a weekly schedule with columns for days (Monday to Friday) and rows for time slots (e.g., 8:00-9:00, 9:00-10:00, etc.). The courses are represented by colored blocks within these slots. The courses are organized into blocs, with some courses appearing in multiple weeks. The colors used include yellow, green, orange, red, and blue. The timetables show a variety of course combinations, with some courses appearing in multiple weeks and others appearing only once. The overall structure is consistent across all 20 timetables, indicating a standardized schedule for the semester.

Organisation 2nd semester M1

The image displays a grid of 20 small Gantt charts, organized into 5 rows and 4 columns. Each chart represents a weekly schedule from 12/10 to 12/20. The charts show the timing and duration of various tasks, color-coded as follows:

- Task 1 (Yellow):** Typically spans from 12/10 to 12/12.
- Task 2 (Blue):** Typically spans from 12/13 to 12/15.
- Task 3 (Green):** Typically spans from 12/16 to 12/18.
- Task 4 (Red):** Typically spans from 12/19 to 12/20.

Each chart includes a vertical axis with dates (12/10, 12/11, 12/12, 12/13, 12/14, 12/15, 12/16, 12/17, 12/18, 12/19, 12/20) and a horizontal axis with days of the week (Sun, Mon, Tue, Wed, Thu, Fri, Sat, Sun). The tasks are represented by colored bars with labels such as 'Task 1', 'Task 2', 'Task 3', and 'Task 4'.

Organisation 1st semester M2

The semester is organised in bloc courses:

| | | | | |
|--|--|--|--|--|
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

Check on the ENT!

First semester - M1

- UE 701 Probabilities, Statistics and Algorithms
- UE 702 Design and acquisition of corpora
- UE 703 Software Engineering
- UE 704 Linguistics for NLP
- UE 705 Project and Language

- UE 701 Probabilities, Statistics and Algorithms
 - EC1 Probabilities and statistics L. Burhy / Marianne Clausel
 - EC2 Algorithms M. Amblard
- UE 702 Design and acquisition of corpora

- UE 703 Software Engineering

- UE 704 Linguistics for NLP

- UE 705 Project and Language

- UE 701 Probabilities, Statistics and Algorithms
 - EC1 Probabilities and statistics L. Burhy / Marianne Clausel
 - EC2 Algorithms M. Amblard
- UE 702 Design and acquisition of corpora
 - EC1 Written corpora B. Guillaume
 - EC2 Logic P. de Groot
- UE 703 Software Engineering

- UE 704 Linguistics for NLP

- UE 705 Project and Language

- UE 701 Probabilities, Statistics and Algorithms
 - EC1 Probabilities and statistics L. Burhy / Marianne Clausel
 - EC2 Algorithms M. Amblard
- UE 702 Design and acquisition of corpora
 - EC1 Written corpora B. Guillaume
 - EC2 Logic P. de Groote
- UE 703 Software Engineering
 - EC1 Software design and development, Functional analysis and specifications, Project management
A. Ducreau, P. Nourissier, R. Altmeyer, Eric François
- UE 704 Linguistics for NLP

- UE 705 Project and Language

- UE 701 Probabilities, Statistics and Algorithms
 - EC1 Probabilities and statistics L. Burhy / Marianne Clausel
 - EC2 Algorithms M. Amblard
- UE 702 Design and acquisition of corpora
 - EC1 Written corpora B. Guillaume
 - EC2 Logic P. de Groote
- UE 703 Software Engineering
 - EC1 Software design and development, Functional analysis and specifications, Project management
A. Ducreau, P. Nourissier, R. Altmeyer, Eric François
- UE 704 Linguistics for NLP
 - EC2 Methods for Natural Language Processing M. Amblard
 - UE702 EC1 Phonology K. Bartkova
 - UE702 EC1 Morphology F. Nammer
- UE 705 Project and Language

- UE 701 Probabilities, Statistics and Algorithms
 - EC1 Probabilities and statistics L. Burhy / Marianne Clausel
 - EC2 Algorithms M. Amblard
- UE 702 Design and acquisition of corpora
 - EC1 Written corpora B. Guillaume
 - EC2 Logic P. de Groote
- UE 703 Software Engineering
 - EC1 Software design and development, Functional analysis and specifications, Project management
A. Ducreau, P. Nourissier, R. Altmeyer, Eric François
- UE 704 Linguistics for NLP
 - EC2 Methods for Natural Language Processing M. Amblard
 - UE702 EC1 Phonology K. Bartkova
 - UE702 EC1 Morphology F. Nammer
- UE 705 Project and Language
 - EC1 Hands-on Project M. Amblard
 - EC2-en english C. Martin
 - EC2-fr français J. Mendes Antunes

- One special organisation for UE 701

Important

- One special organisation for **UE 701**

- **Rooms** 103-104
- Courses with SC
- **Campus!** CLSH

Second semester - M1

- UE 801 Machine learning and semantic web
- UE 802 Formal tools
- UE 803 Data Processing
- UE 804 Linguistics for NLP
- UE 805 Project and Language

- UE 801 Machine learning and semantic web
 - EC1 Machine learning theory
 - EC2 Semantic web
- UE 802 Formal tools

M. Clausel
M. Couceiro

- UE 803 Data Processing
- UE 804 Linguistics for NLP
- UE 805 Project and Language

- UE 801 Machine learning and semantic web
 - EC1 Machine learning theory M. Clausel
 - EC2 Semantic web M. Couceiro
- UE 802 Formal tools
 - EC1 Formal languages P. de Groot
 - EC2 Speech corpora Y. Laprie
 - EC3 Calculability and complexity M. Couceiro
- UE 803 Data Processing
- UE 804 Linguistics for NLP
- UE 805 Project and Language

- UE 801 Machine learning and semantic web
 - EC1 Machine learning theory M. Clausel
 - EC2 Semantic web M. Couceiro
- UE 802 Formal tools
 - EC1 Formal languages P. de Groot
 - EC2 Speech corpora Y. Laprie
 - EC3 Calculability and complexity M. Couceiro
- UE 803 Data Processing
 - EC1 Storage and retrieval C. Gardent
 - EC2 Data analysis Y. Parmentier
- UE 804 Linguistics for NLP

- UE 805 Project and Language

- UE 801 Machine learning and semantic web
 - EC1 Machine learning theory M. Clausel
 - EC2 Semantic web M. Couceiro
- UE 802 Formal tools
 - EC1 Formal languages P. de Groot
 - EC2 Speech corpora Y. Laprie
 - EC3 Calculability and complexity M. Couceiro
- UE 803 Data Processing
 - EC1 Storage and retrieval C. Gardent
 - EC2 Data analysis Y. Parmentier
- UE 804 Linguistics for NLP
 - EC1 Lexicology : lexical units and phraseology A. Polguere
 - UE802 EC1 Syntax A. Polguere
 - UE802 EC1 Semantics I. Comorovski
- UE 805 Project and Language

- UE 801 Machine learning and semantic web
 - EC1 Machine learning theory M. Clausel
 - EC2 Semantic web M. Couceiro
- UE 802 Formal tools
 - EC1 Formal languages P. de Groot
 - EC2 Speech corpora Y. Laprie
 - EC3 Calculability and complexity M. Couceiro
- UE 803 Data Processing
 - EC1 Storage and retrieval C. Gardent
 - EC2 Data analysis Y. Parmentier
- UE 804 Linguistics for NLP
 - EC1 Lexicology : lexical units and phraseology A. Polguere
 - UE802 EC1 Syntax A. Polguere
 - UE802 EC1 Semantics I. Comorovski
- UE 805 Project and Language
 - EC1 Supervised Project M. Amblard
 - EC2-en english C. Martin
 - EC2-fr français J. Mendes Antunes

See the website!

Supervised Project

- a work in pairs or trinomials
- in parallel of the courses
- require a regular follow-up by the tutors
- estimated at a working day per group and per week
- first part consists of a bibliographical work which leads to on the production of a written report to be submitted at the end of semester 7 ;
- the second part, organized in continuity, must include a part realization (survey, data study, programming...), and gives rise to a written report and a defense.
- Important to work on the supervised project as soon as possible

- September 29th, 14h-17h, presentation of topics to students
- October 1st: choice of students / group
- October 2nd: assignment of projects
- October 5th: beginning of the bibliography part
- December 18th: dead-line for the bibliography report
- January 18th 2021: start of the realisation part
- June 18th: dead-line final report
- June 28 and/or 29: defenses

First semester - M2

- UE 901 Deep Learning and Data Mining
- UE 902 Text and speech processing
- UE 903 NLP and Discourse

- UE 901 Deep Learning and Data Mining
 - EC1 Neural Networks, Deep Neural Networks E. Vincent
 - EC2 Data mining (structured data and text) Y. Toussaint+A. Brun
 - EC3 Collaborative filtering A. Boyer + A. Brun
- UE 902 Text and speech processing

- UE 903 NLP and Discourse

- UE 901 Deep Learning and Data Mining
 - EC1 Neural Networks, Deep Neural Networks E. Vincent
 - EC2 Data mining (structured data and text) Y. Toussaint+A. Brun
 - EC3 Collaborative filtering A. Boyer + A. Brun
- UE 902 Text and speech processing
 - EC1 Speech processing R.Serizel
 - EC2 Processing Textual Data K. Smaili
 - EC3 Terminology and ontology M. Constant + M. Couceiro
- UE 903 NLP and Discourse

- UE 901 Deep Learning and Data Mining
 - EC1 Neural Networks, Deep Neural Networks E. Vincent
 - EC2 Data mining (structured data and text) Y. Toussaint+A. Brun
 - EC3 Collaborative filtering A. Boyer + A. Brun
- UE 902 Text and speech processing
 - EC1 Speech processing R.Serizel
 - EC2 Processing Textual Data K. Smaili
 - EC3 Terminology and ontology M. Constant + M. Couceiro
- UE 903 NLP and Discourse
 - EC1 Application to text C. Gardent + D. Langlois
 - EC2 Computational semantics P. de Groote
 - EC3 Discourse and Dialog modeling M. Amblard + Chloé Braud

- UE 904 Lexicons and Grammars for NLP
- UE 905 Project and Language

- UE 904 Lexicons and Grammars for NLP
 - EC1 Diachronic and synchronic lexicology Y. Keromnes
 - EC2 Lexical resources C. Cerisara + Thimothee Mickus
 - EC3 Syntactic framework M. Amblard
- UE 905 Project and Language

- UE 904 Lexicons and Grammars for NLP
 - EC1 Diachronic and synchronic lexicology Y. Keromnes
 - EC2 Lexical resources C. Cerisara + Thimothee Mickus
 - EC3 Syntactic framework M. Amblard
- UE 905 Project and Language
 - EC1 Software Project M. Couceiro + Ajinka Kulkarmi
 - EC2 Law and ethics
 - EC3a Research methods M. Boritchev
 - EC3a Professional integration
 - EC4-en english C. Martin
 - EC2-fr français J. Mendes Antunes

Important

- Rooms 103-104
- Courses with SC
- Campus! CLSH

See the website!

Second semester - M2

Master thesis, 30 ects

- Based on a research work or professional internship
- 5 months between February and September
- Defence in the course of (June or) September
- A selection of topics from Loria will be offered

Master thesis, 30 ects

- Based on a research work or professional internship
- 5 months between February and September
- Defence in the course of (June or) September
- A selection of topics from Loria will be offered

Note that : even if you follow 905 EC3a or b there is no strict relation with the localisation of your internship.

Moreover, the research teams choose to welcome or not students .

Start searching for your internship now!

Master Evaluation

- in June or September
- **Report**: 40 to 50 pages + bibliography
Remind that you must present your work and show that you have a personal point of view on a specific issue.
- 40 minutes of **defence**: 20 mins of presentation (slideshow) + 20 minutes (at least) of questions from the jury
- a short report from **supervisors**

See the website!

Study regulation

First semester

- for evaluation, the average is computed from the marks obtained in the all courses with a coefficient proportional to the ects credits
- the semester is passed if the average is at least 10/20 and if no mark is less than 6
- for the students who do not pass the semester after the first session of exam, there is a second session. The marks of the second session automatically replace the corresponding marks of the first session.
- only students who fail the 1st session can attend the 2nd session.

Life

Nancy in September (and later)

- 09.04 – 10.04 Été indien à l'autre Canal (mostly free if you are member)
- 09.11-12-13 l'Enfer (hand-made printing festival)
- 09.11-20 Le livre sur la place
- 10.03-17 Nancy Jazz Pulsation (NJP)
- La Manufacture ,salle Poirel, l'Autre canal, CCAM
- Médiathèque
- Cinema: Caméo, UGC, Kinopolis
- musée aquarium, musée des beaux arts
- My Monkey Gallery, Ergastule, Centre Culturel George Pompidou, It Up, Galerie Lillebone

Plagiarism is BAD!

The last One more thing!

Nancy is easy to bike:

- Atelier Dynamo
- Atelier Vélo de Maxéville