Technical annex – EvAAL Competition (2013)

Version 1.0
2012/11/10

Important: this version of the annex will be refined with the feedback of the competitors. Refined versions will be timely distributed to the competitors by means of the contest@evaal.aaloa.org mailing list.

Technologies

Each team should implement an **activity recognition system (ARS)** to cover the area of the Living Lab (**indoors**). There is no limitation on the number of devices that can be used.

The ARS can be based on a variety of sensors and technologies, including: accelerometers, gyroscopes, magnetometers, pressure sensors, microphones, sensor networks, cameras, mobile phones, etc. The proposed systems may also include combinations of different technologies. Other technologies can be accepted provided they are compatible with the constraints of the hosting living lab. To this purpose competitors wishing to check such compatibility may inquire with the organizers by e-mail (info@evaal.aaloa.org).

The teams should consider possible restrictions related to the availability of power plugs, cable displacement, attachment of devices to walls/furniture in the Living Lab, etc. The requirements of the proposed ARS should be communicated at an early stage in order to make the necessary on-site arrangements. The **Technical Program Committee (TPC)** may exclude an ARS if their deployment is incompatible with the living lab constraints.

Activities that the competitors must recognize

We will classify the activities in the next categories:
- Walking, standing up, sitting down, lying down, bending, falling and cycling (using a stationary bike).
- NULL class which corresponds to any time slices in which activities that we are not interested in are being performed (no relevant action is performed).

Transitions between activities will not be evaluated so "standing up" includes standing still and "sitting down" also includes be seated.
Bending activity includes the positions showed on the image:
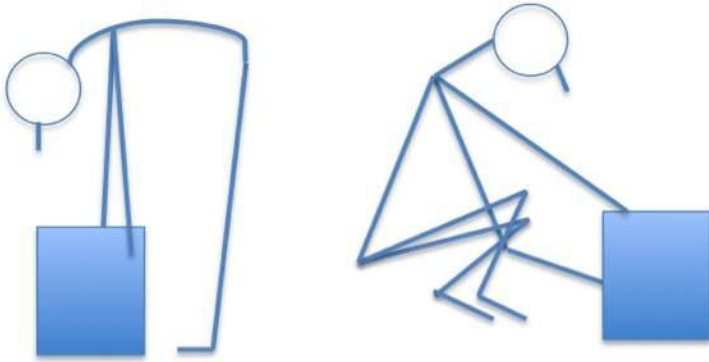
Activities such as running and others that are not evaluated are in Null class.

Duration of activities: Every activity will have duration from 2 seconds to 3 minutes, except the falling activity (0,5 - 3 seconds).

Fall activity includes the next kind of fall [1]:

- Backward fall (both legs straight or with knee flexion).
- Forward fall (On the knees and with or without arm protection).
- Lateral fall.
- Vertical slipping against a wall finishing in sitting position.

Evaluation criteria

In order to evaluate the competing ARS, the TPC will apply the evaluation criteria listed in this document. For each criterion, a numerical score will be awarded. Where possible the score will be measured by direct observation or technical measurement. Whether this is not possible, the score will be determined by the **Evaluation Committee (EC)**. The EC will be composed of some volunteer members of the Technical Program Committee TPC, and will be present during the competition at the Living Lab.

The evaluation criteria are:

1. *Accuracy* – F-measure (2*precision*recall/precision+recall) will be used to measure and compare the accuracy of the ARS.
2. *User acceptance* – Captures how much invasive the ARS is in the user's daily life and thereby the impact perceived by the user; this parameter will be evaluated by the EC.
3. *Recognition delay* – Elapsed time between the instant in which the user begins an activity and the time in which the system recognizes it.
4. *Installation complexity* – A measure of the effort required to install the ARS in a flat/user, measured by the evaluation committee as a function of the person-minutes of work needed to complete the installation (The person-minutes of the first installer will be fully included; the person-minutes of any other installer will be divided by 2).
5. *Interoperability with AAL systems* – Metrics used are: use of open source solutions, use of standards, availability of libraries for development, integration with standard protocols.

Considerations for wearable devices: The recognition system will be used uniquely to that purpose (if the system is a mobile phone, it would not be used to make calls or other tasks). In

any case, none of the wearable devices or devices deployed in the house will be removed/replaced/moved during all the experiments.

The following table presents the overall scoring criteria. Each criterion has a maximum of 10 points, and these are awarded as whole numbers only (i.e., no half points). The weightings shown will be applied to the individual scores in order to determine the overall score:

Evaluation criteria	Weight	Reference values
Accuracy	0.35	F-Measure*10
*User acceptance**	0.2	Score from 0 to 10 given by the evaluation committee.

Recognition delay (RD)	0.2	An integer in the range 0-10.
		
		Score=
		  If RD<=1sec. Score =10
		  If RD>=21sec. score =0
		  If (1< RD < 21)
		      Score = 10- INT((RD-1)/2)

Installation complexity	0.15	Score from 0 to 10 depending on the time required to install the system. The total time is computed as the sum of the minute/person of the first installer (the one that works more) plus the men/minutes of any additional installer divided by 2. If PM is the person/minutes computed in this way, we have:
		
		Score =
		  If PM<=10m. Score =10
		  If PM>=60m. score =0
		  If (10< PM < 60)
		      Score = 10- INT((PM-10)/5)

Interoperability with AAL systems	0.1	Score from 0 to 10, composed of: 2 points for availability of libraries for integration; 2 points for using open solutions for libraries: 2 points for the use of standards: 2 points for availability of tools for testing/monitoring the system; 1 point for availability of sample applications; 1 point for availability of documentation (beyond the submitted paper)

* Some more clarification about the "user acceptance" criteria: The evaluators will judge based mainly on the following criteria, in no particular order of significance:
- Ease of wearing on-body sensors and of taking them off
- Ease of keeping on-body sensors during normal domestic activities
- The user must start the system or it is always logging the activities being recognized
- Batteries recharge rate of the ARS (if needed)

Examples of questions to which evaluators will answer about a wearable device are:
- Does a wearable device exist?
- Does one feel something on his body?
- Is it big?
- Is it washable?

- How unobtrusive is it?
- Can it be easily lost?
- Can it be safely taken outside?
- Does one feel "observed" all the time for wearing it?
- Does one have to configure it?
- Is it easy to notice if it is broken or malfunctioning?
- Can it be completely hidden?

Benchmark Testing

The score for measurable criteria for each competing system will be evaluated by means of benchmark tests (prepared by the organizing committee). For this purpose each team will be allocated a precise time slot at the living lab, during which the benchmark tests will be carried out. The benchmark consists of a set of tests, each of which will contribute to the scores in the assessment of the system. The EC will ensure that the benchmark tests are applied correctly to each system. The evaluation process will also assign scores to the system for the criteria that cannot be assessed directly through benchmark testing.

When both benchmark testing (criteria 1, 3 and 4) and the evaluation by the EC (criteria 2 and 5) have been completed, the overall score for each system will be calculated using the weightings shown above. All final scores will be disclosed at the end of the competition, and the systems ranked according to this final score.

The time slot for benchmark testing is divided in three parts:
- In the first part, the competing team will deploy and configure their system in the living lab and on will equip and instruct the actor playing the role of a typical user. This part should last no more than *60* minutes. This parameter can be re-defined in the next revisions of this document.
- In the second part the benchmark will be applied. During this phase the competitors will have the opportunity to perform only short reconfigurations of their systems (anyway removal, replacement, movement of fixed devices or addition of new devices will not be allowed). In any case, this part should be concluded in *20* minutes (strict deadline). This parameter can be re-defined in the next revisions of this document. The performances used to evaluate the ARS is defined in [this section](#).
- In the last part the teams will remove the system from the living lab and the actor in order to enable the installation of the next competing system (also this part will have strict deadlines).

Competing teams who will fail to meet the deadlines in parts 1 and 3 will be given the minimum score for each criterion related to the benchmark test. Furthermore, systems should be kept active and working during all of the second part. If benchmark testing in the second part is not completed, the system will be awarded a minimum score for all the missing tests.

Actor performance

During the second part, the ARS will be evaluated. An actor (an EC member or a volunteer), will perform a predefined physical activity trip across the smart home. Audio signals will be used to synchronize the actor movements in each performance in order to get the same ground truth for all the participants. For instance, when the actor must change his performance to "walking" activity he will hear the next audio signal: "Walking in 3, 2, 1, now!" When the word "now" is said another EC member will push a radio button to identify the end of the previous activity he will push it again once the transition is finished and the activity is being performed. For example, if the actor is standing still and hears "Cycling in 3, 2, 1, now!" the transition could last some seconds, so until the actor do not pedal, the button is not pressed the second time.

The path followed by the actor and the activities performed will be the same for each test, and it will not be disclosed to competitors before the application of the benchmarks. Notice the smart home has an indoor garden and the actor could also perform activities there.

The environment will be made as much as possible similar to a house. This means that, if possible, there will typical appliances on, neighbour's WiFi AP on, cellular phones on, etc. The kind of environmental noise will be defined during the induction phase of the competition.

There will be two performances:
- The first one performed by a young actor (<65) that will try to simulate an elderly person. Different kinds of falls are allowed here.
- The second one performed by an elderly person (>65) who will not fall at all. If we cannot obtain a volunteer of this age to be present during all the competition, the first actor will substitute him/her.

In order to evaluate the accuracy of the competing systems, the organizers will compare the output of the systems with the ground truth: a timestamp labelled data. The accuracy evaluation is clarified in this section

Details will be made public to the competitors as soon as the decisions in merit will be taken by the TPC.

Accuracy evaluation

F-measure=2*Average_Precision*Average_Recall/(Average_Precision+Average_recall) is used to evaluate the accuracy. Due to F-measure is based on instances, a 250-milisecond time-slot will be used to define it. If the team recognizes correctly the activity at the **end** of the window (after subtracting the algorithm delay) the instance retrieved will be relevant.
In the next figure a toy example is shown with only 4 activities and 1-second time-slots. As you can see the second instance (red time-window) is well classified due to inferred label at the end of the time-slot is A (no matter about the initial D-label). So if many samples are produced in a time slot the only valid sample will be the last one.
If a time-slot window finish in a transition, only the non-transition part of the instance will be evaluated (it will last less than 1 second).
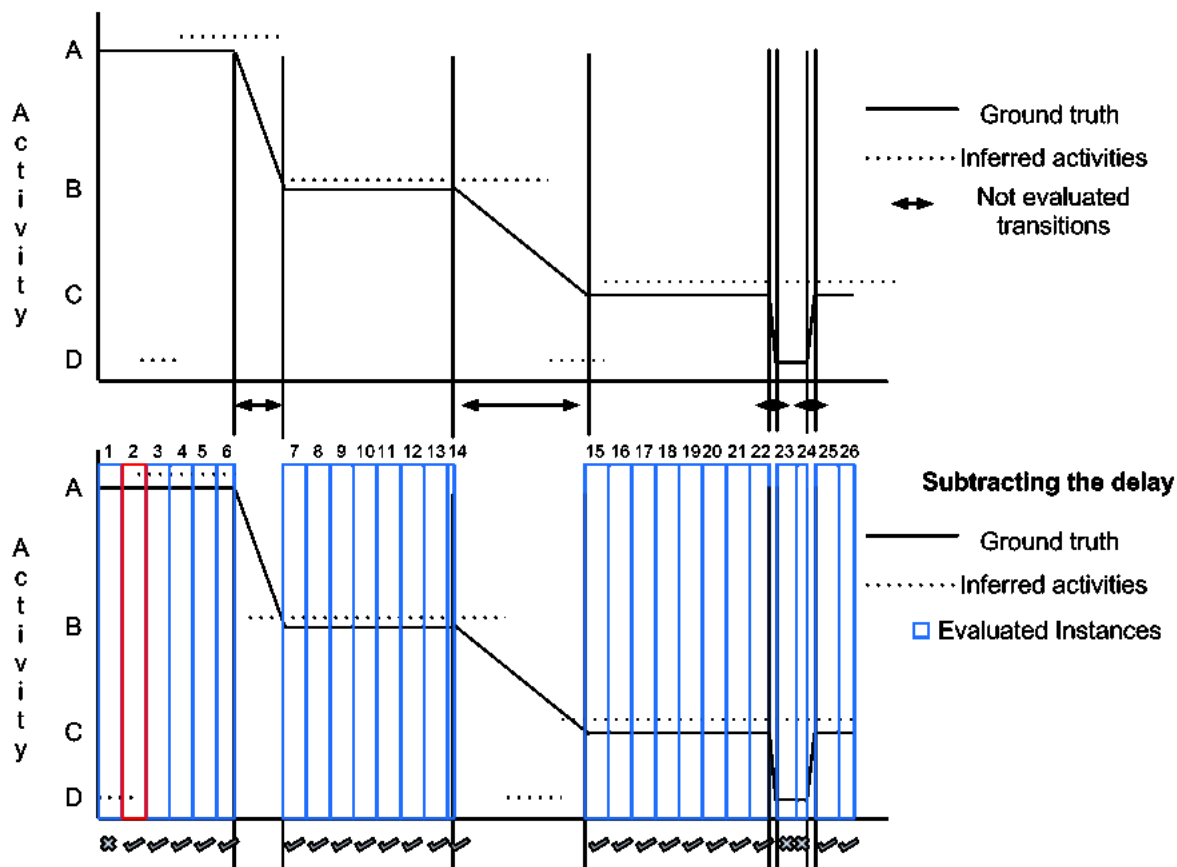
	A	B	C	D	Average
Recall	5/6=0,833	8/8	10/10	0/2	(0,833+1+1+0)/4=0,7083
Precision	5/5	8/8	10/12=0,833	0/1	(1+1+0,833+0)/4=0,7083

F-Measure	0,7083

Our intention is to make the evaluation on line, but we will be logging all events, so we will be able to repeat (and confirm) offline this evaluation. This also in case in which, for any reason, the online evaluation does not work.

Devices and interconnection:

You are expected to bring your own sensors and devices. If your sensors are connected to a laptop, you must connect your laptop with our server to access, in real time, to the data that your activity recognition system produces. We will assist the interconnection between your laptop and our server; in particular, we will deliver software that will enable this. You will have to connect to this software, and we will give you full support for this.

If you need something specific we will ask the responsible of the living lab, our intention is to provide as much support as we can to competitors.

Last year competition

In order to clarify the explained concepts, here you can see a video about last year performance: http://vimeo.com/52843550

As you can see the performance was made by a young actor, acting as himself. **This year the actor we will try to simulate an elderly person.**

The ground truth associated to this video is the next:

Time slots (500 ms)
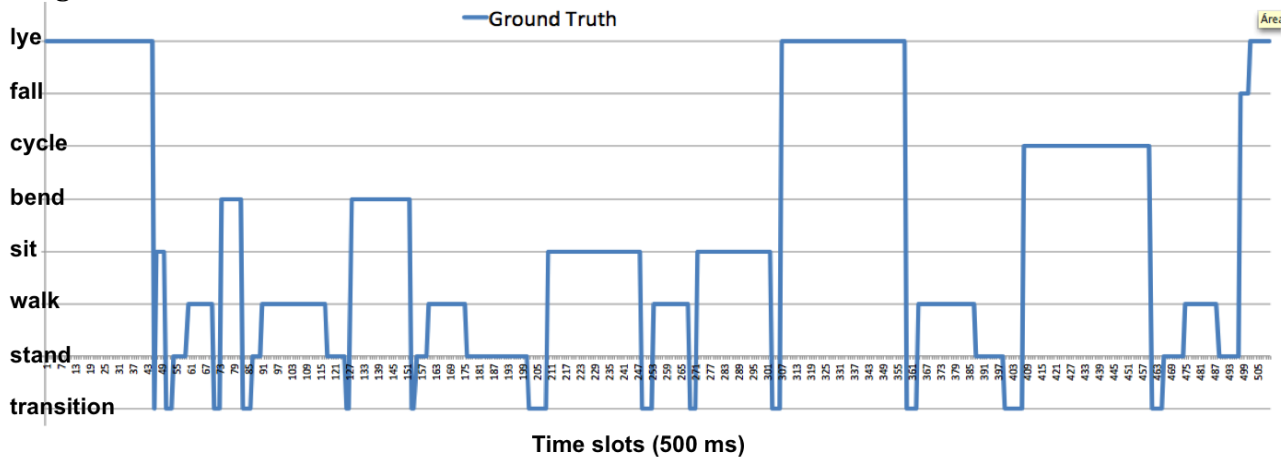
As you can observe, some transitions (all except stand->walk, walk->stand, stand->fall, fall->lye) are labelled as "transition". This label was not evaluated. **This year we will continue without evaluating it.**

After the actor performance finishes, the Accuracy criteria is calculated subtracting time slots (500, 1000, 1500, ... 30000ms) each time in order to get the best Accuracy value. Once obtained the best Accuracy mark we obtain the Recognition Delay (the time subtracted). **Last year we used 500 ms time-slots, this year we will use 250 ms time-slots.**

Here you can see an example of accuracy (after subtract the Recognition Delay) calculation: https://www.dropbox.com/s/n8ixmfehy6sm578/2012%20Accuracy%20Evaluation.xlsx

Bibliography

[1] N. Noury, A. Fleury, P. Rumeau, A. K. Bourke, G. ÓLaighin, V. Rialle, and J. Lundy, "Fall detection - principles and methods," in Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS 2007), 2007, pp. 1663–1666.

Environment

Definition

The CIAmI Living Lab is an approximately 90 m² infrastructure that simulates the real environment of a citizen´s home combined with the existence of Information and Communications Technologies (ICT) massively distributed across the physical space, but as much invisible as possible to the people living in it.
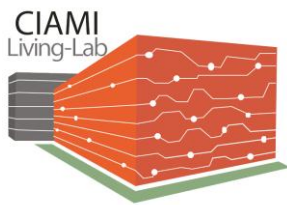

Figure 1 - CIAmI Living Lab logotipe

Location

- CIAmI Living Lab is placed on the Valencia´s Industrial Park (Parque Tecnológico de Valencia, Paterna) on Spain.
- CIAmI Living Lab has been built inside the private company Soluciones Tecnológicas para la Salud y el Bienestar industrial bay.
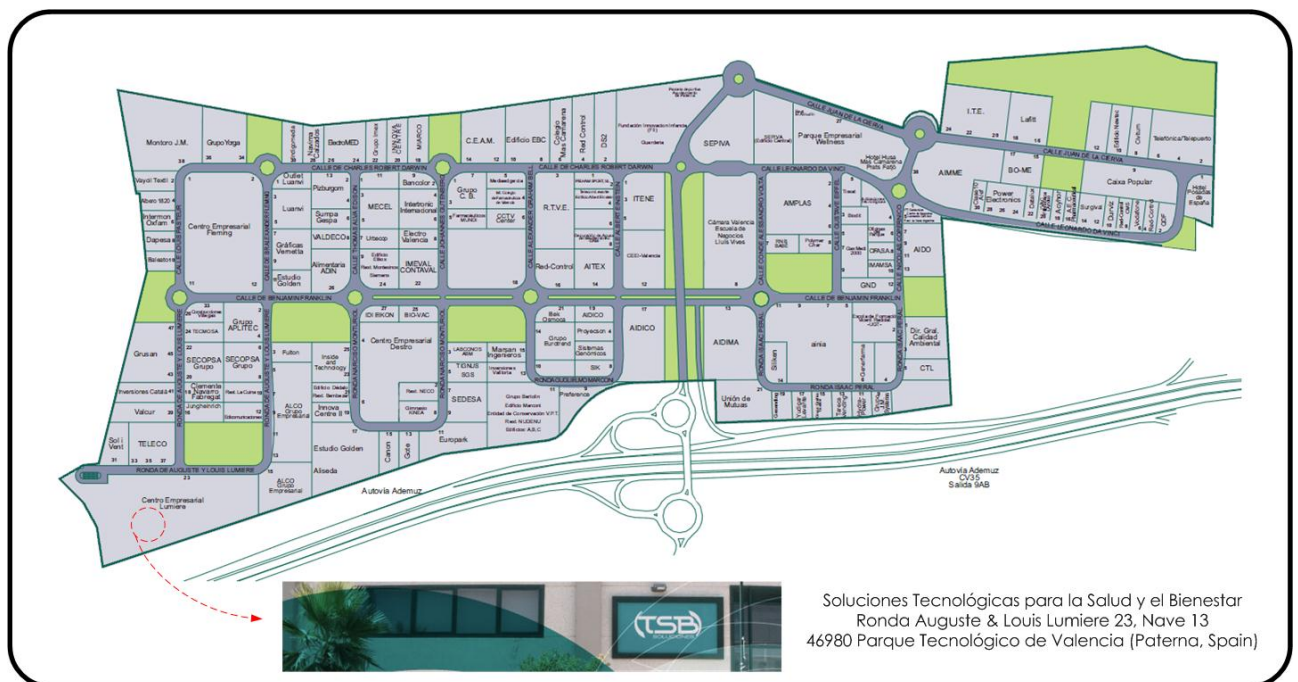

Figure 2 - CIAmI Living Lab location

Infrastructure

- CIAmI Living Lab occupies the ground floor (level sub-1) on the building (please, find attached the document "**Anex2evAALCompetitionCIAmILivingLabPlans.pdf**" where you can see the CIAmI Living Lab plans in high resolution).
- Next to the CIAmI Living Lab there are some laboratories and devices working with RF wireless technologies at 2.4 GHz.
- Near to the CIAmI Living Lab there is a server room with servers and communications equipments (WiFi routers, switches, etc) in 7/24 operation mode.
- The EvAAL competition candidates can install its indoor ARS only in the area reserved for the challenge (please, see the figure number 3).
- CIAmI Living Lab is composed by a control room, a 3D immersive room, a bathroom, a kitchen, a living room and a bedroom. The indoor localization system could be installed in either.
- CIAmI Living Lab has been designed with an open distribution (like a loft). It's very important to note that only the control room, the 3D immersive room and the bathroom are closed rooms.
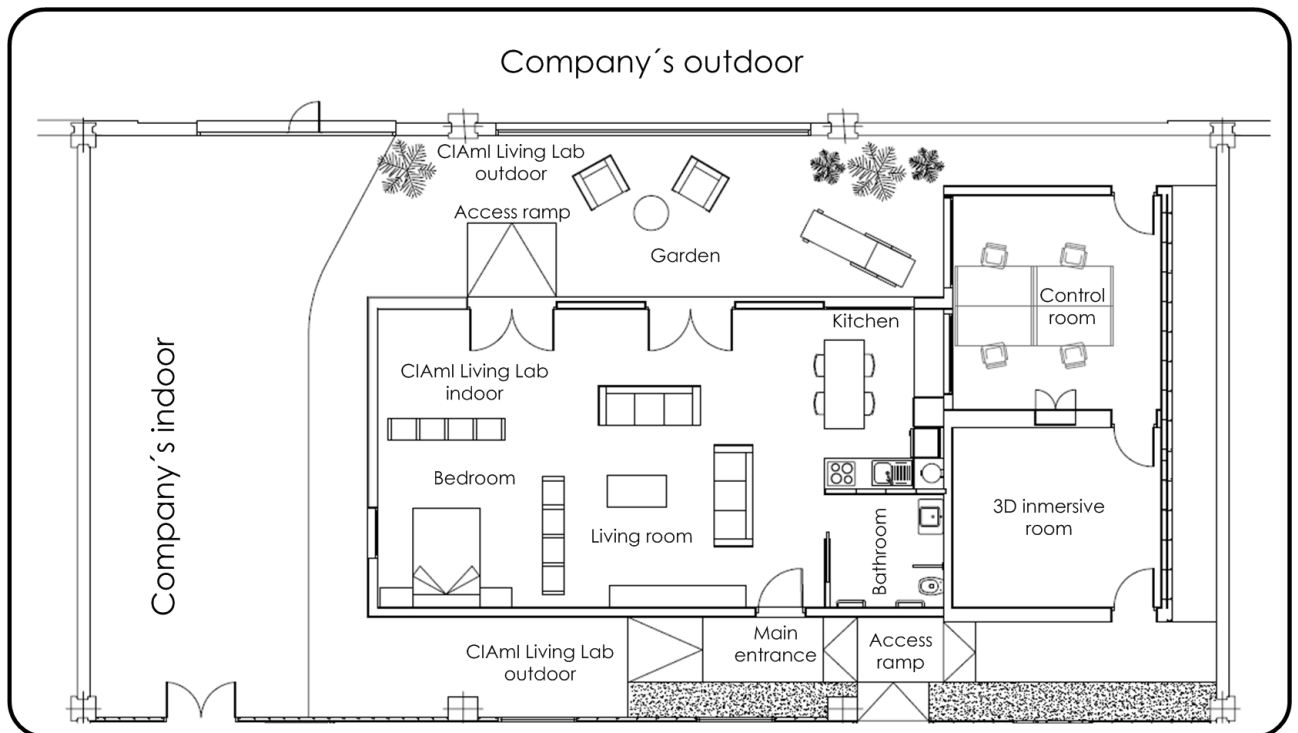

Figure 3 – Restricted area to the EvAAL Competition

Architectural requirements and materials:

- The industrial bay's outdoor walls are reinforced concrete.
- The industrial bay's indoor walls and the CIAmI Living Lab's walls are dry walls or prefabricated modules of plasterboard (Pladur®).

- Easily removable floor (600 x 600 mm width and 40 mm thickness) wood laminated.
- Easily removable plaster ceiling (please, find attached the document "**Anex2evAALCompetitionCIAmILivingLabPlans.pdf**" where you can see the CIAmI Living Lab plans in high resolution).
- Bathroom walls and floor are covered with ceramics tiles, and floor is not removable, only a small part.

Removable floor   Removable ceiling

Figure 4 - Removable floor and ceiling modules at CIAmI Living Lab

- It´s allowed installing devices on the floor, the ceiling and the CIAmI Living Lab walls without drilling in it (double-sided tapes are allowed for temporary installations).
- The installation above the removable ceiling in the 3D immersive room is more difficult and complex than in the other rooms because the distance between the floor and the removable ceiling is higher due to the 3D immersive environment restrictions (please, find attached the document "**Anex2evAALCompetitionCIAmILivingLabPlans.pdf**" where you can see the CIAmI Living Lab plans in high resolution).
- There are only two traditional doors in the CIAmI Living Lab manufactured with wood: the main entrance to the laboratory and the bathroom´s door.
- The control room and the kitchen are separated by a unidirectional vision glass.
- The light in the control and the 3D immersive rooms are fluorescent tubes but in the other rooms are low consumption halogen lamps.
- There is direct communication between the control room and the other rooms (except in the 3D immersive rooms) therefore passing wires is simple and fast.
- Thanks to the unidirectional glass vision between the control room and the kitchen there is direct vision in the entire laboratory (except in the bathroom for ethical reasons).
- CIAmI Living Lab is a laboratory with home look where anyone could live with total comfort and safety.

Figure 5 - Indoor and outdoor views of the CIAmI Living Lab

Communications requirements

- Ethernet sockets and WiFi capabilities across the CIAmI Living Lab.
- High number of power sockets around the CIAmI Living Lab.
- Appliances with home automation capabilities (Maior-Domo® by Fagor).
- Heating, Ventilating and Air Conditioning (HVAC) system.
- Devices could be installed hidden to the user´s senses.
- Oversized wiring pipes and electric boxes (inside the walls).

Accessibility requirements

- Universal design applied in all areas of home construction.
- In compliance with architectural accessibility rules.
- Accesible ramps at all entrances of the CIAmI Living Lab.
- High reconfigurability of all elements in the laboratory.
- Wide corridors (the spaces have been designed to allow anyone pass including those people with they are in wheelchairs or with other assistive products).

- Interaction elements of the laboratory have been signalized and organized to minimize risks and errors.
- Handle openers require a minimum effort to be used.
- Pushers and switchers have enough size so they can be used with fingers, the fist or other parts of the body and they have been situated in places with enough space and at a properly height so they can be reached by any user, independently if the user is sitting or standing.

Available technologies

- Software and hardware development tools for several technologies like Bluetooth, NFC or MHP for instance.
- Full list of home automation system´s sensors and actuators (KNX, X10 and Zigbee Home Automation Profile).
- Media Center with high performance connected to a 37 inches television.
- IP radio with DLNA capabilities.
- Digital television set top boxes with IP and MHP capabilities.
- Smartphones and tabletsPC with Android OS.
- IP cameras.
- Digital frame with NFC support.
- Environmental microphones with high performance.
- Wearable smart sensors.
- Several panel and touch PCs distributed around the laboratory.
- Two 3D auto stereoscopic screens.
- User-friendly and intuitive interaction devices integrated with the environment (for instance, some Nabaztag rabbits and a Roomba vacuum cleaner).
- Neurofeedback equipment for mind peak performance training.
- Inmersive 3D virtual reality system.

More information

- http://www.ciami.es

- If you have some questions or need some additional information about the CIAmI Living Lab facilities you can send an email (please, pay attention that only questions by email will be accepted) to the laboratory´s technical manager (Ángel Martínez, amartinez@tsbtecnologias.es).